Bias and MSE Analysis of the IV Estimator Under Weak Identification with Application to Bias Correction*

John Chao
University of Maryland

Norman R. Swanson
Purdue University

First Draft: January 2000
Current Version: August 2001

Abstract

We provide results on properties of the IV estimator in the presence of weak instruments, beginning with the derivation of analytical formulae for the asymptotic bias (ABIAS) and mean squared error (AMSE). We also obtain approximations for the ABIAS and AMSE formulae based on an asymptotic scheme; which, loosely speaking, 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. One consequence of the asymptotic framework adopted here is that when the sample size and the number of instruments are allowed to approach infinity in a particular sequential manner, we obtain consistent estimators for the ABIAS and AMSE. This in turn suggests a number of bias corrected OLS and IV estimators, which we outline and examine. We also note that, under stronger but more primitive conditions than used in this paper, our bias-corrected estimators can be justified on the basis of a pathwise asymptotic scheme which takes the number of instruments to infinity as a function of the sample size, although this particular result is proved elsewhere. Finally, we note that the bias-corrected IV estimators proposed here are also robust in the sense that they would remain consistent in a conventional asymptotic setup where the model is fully identified. A series of Monte Carlo experiments documents the gains in bias reduction when our bias adjusted estimators are used instead of standard IV and OLS estimators.

JEL classification: C12, C22.

Keywords: confluent hypergeometric function, Laplace approximation, local-to-zero asymptotics, weak instruments.

*John Chao: Department of Economics, University of Maryland, College Park, MD, USA 20742, chao@econ.umd.edu. Norman R. Swanson: Department of Economics, Purdue University, West Lafayette, IN, USA 47907, nswanson@mgmt.pudue.edu.

The authors wish to thank Paul Bekker, Graham Elliott, Roberto Mariano, Carmela Quintos, Tim Vogelsang, Eric Zivot, and seminar participants at Cornell University, the University of California - San Diego, the University of Pennsylvania, the University of Rochester, and the 2000 meetings of the World Congress of the Econometric Society for helpful comments.

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 Hahn and Inoue (2000), Nelson and Startz (1990a), Dufour (1996), Shea (1997), Staiger and Stock (1997), Zivot and Wang (1998), and the references contained therein.[1] 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 noteable 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. To formalize the notion of having weak instruments, we adopt the local-to-zero asymptotic framework of Staiger and Stock (1997) in the context of a simple simultaneous equations setup with a single structural equation and an arbitrary number of available instruments. An important reason for employing the local-to-zero framework is that simulation studies reported in Staiger and Stock (1997) show this framework to yield a very good approximation for the finite sample distribution of the IV estimator when the quality of the available instruments is poor. Our paper, focuses on a general IV estimator, which does not necessarily make use of all available instruments, and we derive explicit analytical formulae for the asymptotic bias and MSE of this estimator under the local-to-zero framework. We find the asymptotic bias and MSE functions of the IV estimator to depend only on the size of a concentration parameter $\mu'\mu$, the number of instruments used k_{21}, and the second moments of the disturbances of the underlying model. Because the analytical formulae for bias and MSE involve confluent hypergeometric functions and are, thus, complicated, we also derive approximations for these formulae based on an expansion which, loosely speaking, requires the first stage F-statistic for testing instrument relevance to converge to a finite (possibly small) positive limit as the number of instruments approaches infinity. Since this expansion is performed on the analytical formulae for the asymptotic bias and MSE, the resulting 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 k_{21}. Our numerical results show that this expansion give good approximations for the ABIAS and AMSE. Moreover, the approximations allow us to make interesting additional observations. For example, 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 forumlae for the asymptotic bias and MSE, derived under the local-to-zero framework, correspond

[1]Related to the weak instrument literature is a literature which examines the implications for statistical inference when the underlying simultaneous equations model is underidentified, in the sense of not satisfying the usual rank condition for identification. Notable contributions to this literature include Phillips (1989), Choi and Phillips (1992), Kitamura (1994), and the references contained therein.

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.

An important consequence of the sequential limit approach which we adopt here is that consistent estimators for the ABIAS and AMSE can be obtained. In particular, the availability of consistent estimators for the bias enables us to construct bias-corrected OLS and IV estimators, which estimate consistently the structural coefficient of the simultaneous equations model even when the available instruments are weak in the local-to-zero sense. Moreover, we note that, under some stronger but more primitive alternative conditions, the bias-corrected estimators proposed here are also consistent under a pathwise asymptotic scheme whereby the number of instruments is taken to approach infinity as a function of the sample size, a result which is proven elsewhere (i.e. see Chao and Swanson (2001)). Additionally, we have shown that in the conventional setup where the model is fully identified, all but one of our proposed bias-corrected estimators remain consistent. A series of Monte Carlo experiments documents the gains in bias reduction when our bias adjusted estimators are used instead of standard IV and OLS estimators.

This rest of the paper is organized as follows. Section 2 contains preliminaries, including the model, assumptions, and notation. Section 3 presents formulae for the ABIAS and AMSE, and discusses some properties of the formulae. Section 4 develops ABIAS and AMSE approximations. Section 5 discusses the consistent estimation of the ABIAS and AMSE, and suggests a number of bias corrected OLS and IV estimators. Section 6 summarizes various numerical calculations used to assess the accuracy of our approximations, and in Section 7, a series of Monte Carlo experiments are used to illustrate the performance of our proposed bias corrected estimators. Concluding remarks are given in Section 8. All proofs and technical details are contained in the appendices.

Before proceeding, we briefly introduce some notation. In the sequel, we use the symbol "$\Longrightarrow$" to denote convergence in distribution. "$\equiv$ " denotes equivalence in distribution. $P_X = X(X'X)^{-1}X'$ is the matrix which projects orthogonally onto the range space of X and $M_X = I - P_X$, and $P_{(Z,X)} = P_X + M_X Z\left(Z'M_X Z\right)^{-1} Z'M_X$ and $M_{(Z,X)} = M_X - M_X Z\left(Z'M_X Z\right)^{-1} Z'M_X$.

2 Setup

Consider the simultaneous equations model (SEM):

$$\begin{aligned} y_1 &= y_2\beta + X\gamma + u, \quad (1)\\ y_2 &= Z\Pi + X\Phi + v, \quad (2) \end{aligned}$$

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[2]. Let u_t and v_t denote the t^{th} component of the random vectors u and v,

[2]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. In addition, as we will point out in Section 5 of this paper, many of the results of that section can, in fact, be generalized in a straightforward manner to a model with multiple endogenous explanatory variables.

respectively; and let Z_t' and X_t' denote the t^{th} row of the matrices Z and X, respectively. Additionally, let $w_t = (u_t, v_t)'$ and let $\overline{Z}_t = (X_t', Z_t')'$; 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 formalize the notion of weak instruments by modeling Π to be a parameter sequence that is local to zero.[3] In particular, we make the following assumption.

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

Also, following Staiger and Stock (1997), we assume that the data generating process of the exogenous variables, $\overline{Z} = (X, Z)$, and of the disturbances, (u, v), is such that the following moment convergence results hold.

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)* $(T^{-1/2}X'u, T^{-1/2}Z'u, T^{-1/2}X'v, T^{-1/2}Z'v) \Longrightarrow (\psi_{Xu}, \psi_{Zu}, \psi_{Xv}, \psi_{Zv})$, *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 the parameter β in equation (1) above, where the IV estimator may not make use of all available instruments. Define the IV estimator as: $\widehat{\beta}_{IV} = (y_2'(P_H - P_X)y_2)^{-1}(y_2'(P_H - P_X)y_1)$, 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 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 $\Pi_T = (\Pi_{1,T}', \Pi_{2,T}')' = (C_1'/\sqrt{T}, C_2'/\sqrt{T})'$, $T^{-\frac{1}{2}}Z'u = (T^{-\frac{1}{2}}(Z_1'u)', T^{-\frac{1}{2}}(Z_2'u)')'$, $T^{-\frac{1}{2}}Z'v = (T^{-\frac{1}{2}}(Z_1'v)', T^{-\frac{1}{2}}(Z_2'v)')'$, $\psi_{Zu} = (\psi_{Z_1u}', \psi_{Z_2u}')'$, and $\psi_{Zv} = (\psi_{Z_1v}', \psi_{Z_2v}')'$, where from part (iii) of Assumption 2 we have that $(T^{-\frac{1}{2}}(Z_1'u)', T^{-\frac{1}{2}}(Z_2'u)', T^{-\frac{1}{2}}(Z_1'v)', T^{-\frac{1}{2}}(Z_2'v)')' \Rightarrow (\psi_{Z_1u}', \psi_{Z_2u}', \psi_{Z_1v}', \psi_{Z_2v}')'$. Furthermore, partition Q conformably with $\overline{Z} = (X, Z_1, Z_2)$ as

$$Q = \begin{pmatrix} Q_{XX} & Q_{XZ_1} & Q_{XZ_2} \\ Q_{Z_1X} & Q_{Z_1Z_1} & Q_{Z_1Z_2} \\ Q_{Z_2X} & Q_{Z_2Z_1} & Q_{Z_2Z_2} \end{pmatrix}. \tag{3}$$

Finally, define

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12}' & \Omega_{22} \end{pmatrix} = \begin{pmatrix} Q_{Z_1Z_1} - Q_{Z_1X}Q_{XX}^{-1}Q_{XZ_1} & Q_{Z_1Z_2} - Q_{Z_1X}Q_{XX}^{-1}Q_{XZ_2} \\ Q_{Z_2Z_1} - Q_{Z_2X}Q_{XX}^{-1}Q_{XZ_1} & Q_{Z_2Z_2} - Q_{Z_2X}Q_{XX}^{-1}Q_{XZ_2} \end{pmatrix} \tag{4}$$

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

Assumption 3: $\sup_T E(|U_T|^{2+\delta}) < \infty$, *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.

Note that 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

[3] Similar types of asymptotic approaches to that used here have been used in a number of contexts in recent years in the econometric literature (see e.g. Elliott (1998) and the references contained therein.)

A. Hence, under Assumption 3, the asymptotic bias and MSE correspond to the bias and MSE implied by the limiting distribution of $\widehat{\beta}_{IV,T}$. Note also that for the special case where $(u_t, v_t)' \sim i.i.d.\ N(0, \Sigma)$, $k_2 \geq 4$ implies Assumption 3, since it is well-known that the IV estimator of β under Gaussianity has finite sample moments which exist up to and including the degree of apparent overidentification, as given by the order condition (e.g. see Sawa (1969)). Throughout this paper, we shall assume $k_{21} \geq 4$ so as to ensure that our results 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 it is well-known that the finite sample distribution of LIML in this case has Cauchy-like tails so that no positive integer moment exist. (See Mariano and McDonald (1979) and Phillips (1984, 1985) for various results documenting the non-existence of moments of the finite sample distribution of LIML.) Consequently, we do not consider the LIML estimator in this paper.

3 Asymptotic Bias and MSE: Formulae and Properties

We begin with two theorems which give explicit analytical formulae for the asymptotic bias and MSE of the *IV* estimator in the case with weak instruments. The theorems also characterize some of the properties of the bias and MSE functions.

Theorem 3.1 (Bias) *Given the SEM described by equations (1) and (2), and under Assumptions 1, 2, and 3, the following results hold for $k_{21} \geq 4$:*
(a)

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

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 $\sigma_{uv} \neq 0$.

Theorem 3.2 (MSE): *Given the SEM described by equations (1) and (2), and under Assumptions 1, 2, and 3, the following results hold for $k_{21} \geq 4$:*
(a)

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

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 and where $\mu'\mu$, ρ, $\Gamma(\cdot)$, and ${}_1F_1(\cdot;\cdot;\cdot)$ are as defined in Theorem 3.1.

(b) For k_{21} fixed, as $\mu'\mu \to \infty$, $m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) \to 0$.
(c) For $\mu'\mu$ fixed, as $k_{21} \to \infty$, $m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) \to \sigma_{uv}^2/\sigma_{vv}^2 = \sigma_{uu}\sigma_{vv}^{-1}\rho^2$.
(d) 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 $\sigma_{uv} \neq 0$.

Remark 3.3: (i) Note that the asymptotic bias and MSE formulae, given by expressions (5) and (6) above, involve confluent hypergeometric functions (see Lebedev (1972) for more detailed discussions of confluent hypergeometric functions). It is well known that confluent hypergeometric functions have infinite series representations (e.g. see Slater (1960), pp.2), so that ${}_1F_1(a;\ b;\ x) = \sum_{j=0}^{\infty} \frac{(a)_j}{(b)j} \frac{x^j}{j!}$, where $(a)_j$ denotes the Pochhammer symbol (i.e. $(a)_j = \Gamma(a+j)/\Gamma(a)$ for integer $j \geq 1$, and $(a)_j = 1$ for $j = 0$). It follows that the expressions for the bias and MSE can be written in infinite series form:

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

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

The main merit of these infinite series representations is that they provide explicit formulae for the ABIAS and AMSE of the IV estimator under weak identification, which can be used in numerical calculations. We have also made extensive use of these representations in deriving the properties of the ABIAS and AMSE reported in the above theorems.

(ii) Part (b) of Theorem 3.2, states that the MSE function for $\widehat{\beta}_{IV,T}$ approaches zero as $\mu'\mu \to \infty$. Note that the case where $\mu'\mu \to \infty$ corresponds roughly to the case where the available instruments are not weak, but are instead fully relevant. In this case, then, Theorem 3.2 part (b) establishes that $\widehat{\beta}_{IV,T}$ converges in a mean squared sense to the true value, β_0. It follows, then, that in this case $\widehat{\beta}_{IV,T}$ is a (weakly) consistent estimator of β, a result which also follows from conventional asymptotic analysis with a fully identified model. Hence, results associated with the standard textbook case of good instruments are a limiting special case of our results.

(iii) Observe that under a condition similar to Assumption 3 above, it can be shown that $\lim_{T\to\infty} E\left(\widehat{\beta}_{OLS,T} - \beta_0\right) = \sigma_{uv}/\sigma_{vv}$ and $\lim_{T\to\infty} E\left(\widehat{\beta}_{OLS,T} - \beta_0\right)^2 = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 = \sigma_{uv}^2/\sigma_{vv}^2$, where $\widehat{\beta}_{OLS,T}$ is the OLS estimator. To see this, let $U_T^* = \widehat{\beta}_{OLS,T} - \beta_0$ and assume that $\sup_{T\geq 5} E[|U_T^*|^{2+\delta}] < \infty$. It is well-known that this assumption holds under Gaussian error assumptions for T adequately large, since the finite sample distribution of the OLS estimator in this case has moments which exist up to the order $T-2$ (see Sawa, 1969, for a more detailed discussion of the existence of moments of the OLS estimator). Now, to proceed with the derivation of $\lim_{T\to\infty} E\left(\widehat{\beta}_{OLS,T} - \beta_0\right)$ and $\lim_{T\to\infty} E\left(\widehat{\beta}_{OLS,T} - \beta_0\right)^2$, we note that Staiger and Stock (1997) have shown that under Assumption 2 above, $\widehat{\beta}_{OLS,T} - \beta_0 \xrightarrow{p} \sigma_{uv}/\sigma_{vv}$, as $T \to \infty$. Thus, it follows directly from Theorem 5.4

of Billingsley (1968) that $\lim_{T\to\infty} E\left(\widehat{\beta}_{OLS,T}-\beta_0\right)=\sigma_{uv}/\sigma_{vv}$ and $\lim_{T\to\infty} E\left(\widehat{\beta}_{OLS,T}-\beta_0\right)^2=E(\sigma_{uv}/\sigma_{vv})^2=\sigma_{uu}\sigma_{vv}^{-1}\rho^2$. Moreover, comparing asymptotic bias and MSE of the OLS estimator given here with the bias result obtained in part (c) of Theorem 3.1 and the MSE result obtained in part (c) of Theorem 3.2, respectively, we see that for fixed $\mu'\mu$, the ABIAS and the AMSE of the IV estimator converge to those of the OLS estimator, as $k_{21}\to\infty$.

(iv) Write the asymptotic bias function of $\widehat{\beta}_{IV}$ as $b_{\widehat{\beta}_{IV}}(\mu'\mu,k_{21})=\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho f(\mu'\mu,k_{21})$, where $f(\mu'\mu,k_{21})=e^{-\frac{\mu'\mu}{2}}\ {}_1F_1\left(\frac{k_{21}}{2}-1;\frac{k_{21}}{2};\frac{\mu'\mu}{2}\right)$. From the proof of part (d) of Theorem 3.1, we see that $0<f(\mu'\mu,k_{21})<1$, for $\mu'\mu\in(0,\infty)$ and for positive integer k_{21} such that the bias function exists. Since $\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho=\sigma_{uv}/\sigma_{vv}$ is simply the asymptotic bias of the OLS estimator, it follows that the bias of the IV estimator given in expression (5) has the same sign as the OLS bias. Additionally, note that $|b_{\widehat{\beta}_{IV}}(\mu'\mu,k_{21})|=|\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho|f(\mu'\mu,k_{21})<|\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho|$. Hence, even when the instruments are weak in the sense of Staiger and Stock (1997), the asymptotic bias of the IV estimator is less in absolute magnitude than that of the OLS estimator for $\mu'\mu\neq 0$ and for finite values of k_{21} such that the bias function exists, 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} such that the bias function exists. Our result, therefore, formalize the intuitive discussions given in Bound, Jaeger, and Baker (1995) and Angrist and Krueger (1995) which 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.

(v) Although the asymptotic bias of $\widehat{\beta}_{IV}$ is less than that of $\widehat{\beta}_{OLS}$ for all positive values of $\mu'\mu$ and for all values of k_{21} for which the bias exists, and although the former only tends to the latter as $k_{21}\to\infty$ for a given $\mu'\mu$, the AMSE of $\widehat{\beta}_{IV}$ with weak instruments may, depending on the size of the concentration parameter ($\mu'\mu$) and the number of instruments used (k_{21}), be either greater or less than the AMSE of $\widehat{\beta}_{OLS}$. To see this, consider the example where $\mu'\mu=0$, and note that, in this case, the infinite series form of $m_{\widehat{\beta}_{IV}}(\mu'\mu,k_{21})$ reduces to:

$$m_{\widehat{\beta}_{IV}}(\mu'\mu=0,k_{21})=\sigma_{uu}\sigma_{vv}^{-1}\rho^2\left[1+\left(\frac{1-\rho^2}{\rho^2}\right)\left(\frac{1}{k_{21}-2}\right)\right], \tag{9}$$

which is greater than the AMSE of the OLS estimator for $\rho^2<1$ and for values of k_{21} for which the AMSE of the IV estimator exists. On the other hand, we have already shown that as $\mu'\mu\to\infty$, the AMSE of the IV estimator approaches zero for k_{21} fixed, so that as $\mu'\mu$ grows the AMSE of the IV estimator will eventually become smaller than that of the OLS estimator.

(vi) It is of interest to compare our results with those obtained in the extensive literature on the finite sample properties of IV estimators, and in particular with the results of Richardson and Wu (1971), who obtained the exact bias and MSE of the $2SLS$ estimator for a fixed instrument/Gaussian model[4]. To proceed with

[4]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).

such a comparison, note first that the SEM given by expressions (1) and (2) can alternatively be written in the reduced form:

$$
\begin{aligned}
y_1 &= Z\Gamma_1 + X\Gamma_2 + \varepsilon_1, && (10)\\
y_2 &= Z\Pi + X\Phi + \varepsilon_2, && (11)
\end{aligned}
$$

where $\Gamma_1 = \Pi\beta$, $\Gamma_2 = \Phi\beta + \gamma$, $\varepsilon_2 = v$, and $\varepsilon_1 = u + v\beta = u + \varepsilon_2\beta$. In the finite sample literature on IV estimators, a Gaussian assumption is often made on the disturbances of this reduced form model; that is, it is often assumed that $(\varepsilon_{1t},\ \varepsilon_{2t})' \equiv i.i.d.N(\mathbf{0}, G)$, where ε_{1t} and ε_{2t} denote the t^{th} component of the $T \times 1$ random vectors ε_1 and ε_2, respectively and where G can be partitioned conformably with $(\varepsilon_{1t},\ \varepsilon_{2t})'$ as $G = \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix}$. Now, consider the case where all available instruments are used (i.e. the case where the IV estimator is simply the $2SLS$ estimator). Then, it follows that $\mu'\mu = \sigma_{vv}^{-1} C'\Omega C = \sigma_{vv}^{-1} C'(Q_{ZZ} - Q_{ZX}Q_{XX}^{-1}Q_{XZ})C$. In addition, 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 earlier 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 the bias formula (5) and the MSE formula (6), we see that:

$$
\begin{aligned}
b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) &= -\frac{g_{22}\beta - g_{12}}{g_{22}} e^{-\frac{\mu'\mu}{2}}\ {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right), \text{ and} && (12)\\
m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) &= \frac{g_{11}g_{22} - g_{12}^2}{g_{22}} \left(\frac{1}{k_{21} - 2}\right)(1 + \overline{\beta}^2) e^{-\frac{\mu'\mu}{2}}\ {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right) \\
&\quad + \left(\frac{k_{21} - 3}{k_{21} - 2}\right)\overline{\beta}^2 e^{-\frac{\mu'\mu}{2}}\ {}_1F_1\left(\frac{k_{21}}{2} - 2; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right), && (13)
\end{aligned}
$$

where $\overline{\beta} = (g_{22}\beta - g_{12})(g_{11}g_{22} - g_{12}^2)^{-\frac{1}{2}}$. Comparing expressions (12) and (13) 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 the 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 (12) and (13) above, whereas the expression $\sigma_{vv}^{-1}\Pi'Z'M_X Z\Pi$ appears in the exact formulae reported in Richardson and Wu (1971). Hence, our bias and MSE results are 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.

4 Approximation Results for the Bias and MSE

The bias and MSE functions given in Theorems 3.1 and 3.2 involve confluent hypergeometric functions, and thus have complicated infinite series representations, as discussed in Remark 3.3(i). In this section, we provide approximations for the bias and MSE based on an expansion which holds as k_{21} and $\mu'\mu$ approach infinity such that the ratio $\frac{\mu'\mu}{k_{21}}$, i.e. the population analogue of the first stage F-statistic, approaches a positive finite limit. More precisely, assume that:

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

We show in Theorem 4.1 below that approximations based on Assumption 4 yield simpler expressions for the bias and MSE than the infinite series representations given in Remark 3.3(i). In addition, motivation for the type of approximations based on Assumption 4 is provided in the remarks following Theorem 4.1; and regression results which indicate that our approximations perform well, even when k_{21} is relatively small, are presented in Section 6.

Theorem 4.1 (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 expressions (5) and (6), respectively, in terms of τ^2 and k_{21} so that:*

$$b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho e^{-\frac{\mu'\mu(\tau^2,k_{21})}{2}} {}_1F_1\left(\frac{k_{21}}{2}-1; \frac{k_{21}}{2}; \frac{\mu'\mu(\tau^2,k_{21})}{2}\right), \tag{14}$$

$$\begin{aligned} m_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu(\tau^2,k_{21})}{2}} \left[\frac{1}{\rho^2}\left(\frac{1}{k_{21}-2}\right) {}_1F_1\left(\frac{k_{21}}{2}-1; \frac{k_{21}}{2}; \frac{\mu'\mu(\tau^2,k_{21})}{2}\right)\right. \\ &\left. + \left(\frac{k_{21}-3}{k_{21}-2}\right) {}_1F_1\left(\frac{k_{21}}{2}-2; \frac{k_{21}}{2}; \frac{\mu'\mu(\tau^2,k_{21})}{2}\right)\right] \end{aligned} \tag{15}$$

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

(a)

$$\begin{aligned} b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left\{\left(\frac{1}{1+\tau^2}\right) - \frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left[2 - 4\left(\frac{1}{1+\tau^2}\right)\right.\right. \\ &\left.\left. +2\left(\frac{1}{1+\tau^2}\right)^2\right]\right\} + O(k_{21}^{-2}) \end{aligned} \tag{16}$$

$$= \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left\{\left(\frac{1}{1+\tau^2}\right) - \frac{2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2\right\} + O(k_{21}^{-2}) \tag{17}$$

(b)

$$\begin{aligned} m_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left\{\left(\frac{1}{1+\tau^2}\right)^2 + \left(\frac{1}{\rho^2}\right)\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\right. \\ &\left. -\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\left[7\left(\frac{1}{1+\tau^2}\right) - 12\left(\frac{1}{1+\tau^2}\right)^2 + 6\left(\frac{1}{1+\tau^2}\right)^3\right]\right\} \\ &+O(k_{21}^{-2}) \end{aligned} \tag{18}$$

$$\begin{aligned} &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left\{\left(\frac{1}{1+\tau^2}\right)^2 + \left(\frac{1-\rho^2}{\rho^2}\right)\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right) + \right. \\ &\left. \left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\left[1 - 7\left(\frac{1}{1+\tau^2}\right) + 12\left(\frac{1}{1+\tau^2}\right)^2 - 6\left(\frac{1}{1+\tau^2}\right)^3\right]\right\} \\ &+O(k_{21}^{-2}) \end{aligned} \tag{19}$$

Remark 4.2: (i) Set

$$\widehat{b}_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left[\left(\frac{1}{1+\tau^2}\right) - \frac{2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2\right]. \tag{20}$$

Recall from Remark 3.3(iii) that the ABIAS of the OLS estimator is given by $b_{\widehat{\beta}_{OLS}} = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho$. It follows that, by taking the ratio of the two, we obtain the relative bias measure:

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

Observe that the lead term of expression (21) is $(1+\tau^2)^{-1} = (1+\mu'\mu/k_{21})^{-1}$. Note 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 (see pages 566 and 575 of Staiger and Stock, 1997). 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 $\mu'\mu$, $k_{21} \to \infty$. [5]

(ii) Now, set

$$\begin{aligned}\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left\{\left(\frac{1}{1+\tau^2}\right)^2 + \left(\frac{1-\rho^2}{\rho^2}\right)\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\right. \\ &\quad \left. + \frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left[1 - 7\left(\frac{1}{1+\tau^2}\right) + 12\left(\frac{1}{1+\tau^2}\right)^2 - 6\left(\frac{1}{1+\tau^2}\right)^3\right]\right\},\end{aligned} \tag{22}$$

Note that the lead term of this 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.

(iii) Given $\widehat{b}_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ and $\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$, we can construct an approximation for the asymptotic variance of the IV estimator under the local-to-zero assumption as follows:

[5] 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 $\frac{\mu'\mu}{k_{21}}$, the population analogue of the first stage F-statistic, 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 $T \to \infty$.

$$
\begin{aligned}
V_{\widehat{\beta}_{IV}}(\tau^2,k_{21}) &= m_{\widehat{\beta}_{IV}}(\tau^2,k_{21}) - (b_{\widehat{\beta}_{IV}}(\tau^2,k_{21}))^2 \qquad (23)\\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left\{\left(\frac{1}{1+\tau^2}\right)^2 + \left(\frac{1-\rho^2}{\rho^2}\right)\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right) + \right.\\
&\left.\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\left[1-7\left(\frac{1}{1+\tau^2}\right)+12\left(\frac{1}{1+\tau^2}\right)^2-6\left(\frac{1}{1+\tau^2}\right)^3\right]\right\}\\
&+O\left(k_{21}^{-2}\right)\\
&-\left(\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left\{\left(\frac{1}{1+\tau^2}\right)-\frac{2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2\right\}+O\left(k_{21}^{-2}\right)\right)^2\\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left[\left(\frac{1-\rho^2}{\rho^2}\right)\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)+\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\left\{1\right.\right.\\
&\left.\left.-3\left(\frac{1}{1+\tau^2}\right)+4\left(\frac{1}{1+\tau^2}\right)^2-2\left(\frac{1}{1+\tau^2}\right)^3\right\}\right]+O\left(k_{21}^{-2}\right)\\
&= \widehat{V}_{\widehat{\beta}_{IV}}(\tau^2,k_{21})+O(k_{21}^{-2}),\ \textit{say.} \qquad (24)
\end{aligned}
$$

A desirable feature of the approximation formulae $\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2,k_{21})$ and $\widehat{V}_{\widehat{\beta}_{IV}}(\tau^2,k_{21})$ is that they are non-negative for $k_{21}\geq 4$. This is shown in the following theorem.

Theorem 4.3: *Given* $\tau^2\in(0,\infty)$ *and* $k_{21}\geq 4$,

(a) $\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2,k_{21})\geq 0$,

(b) $\widehat{V}_{\widehat{\beta}_{IV}}(\tau^2,k_{21})\geq 0$,

where $\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2,k_{21})$ *and* $\widehat{V}_{\widehat{\beta}_{IV}}(\tau^2,k_{21})$ *are defined in expressions (22) and (24), respectively.*

(iv) Note that the results of Theorem 4.1 can be viewed as having been obtained from a sequential limit procedure, whereby we first let the sample size T approach infinity to obtain the ABIAS and AMSE functions under weak identification; subsequently, we take both k_{21} and $\mu'\mu$ to infinity, such that $\frac{\mu'\mu}{k_{21}}=\tau^2+Op\left(k_{21}^{-2}\right)$, to obtain the approximation formulae (16) and (18) for a given τ^2. In quite different contexts, Phillips (1998a,b) and Phillips and Moon (1999) have shown that the use of a sequential limit theory often yields very reasonable approximations when the random quantities of interest involve multiple indices. They also show that under mild strengthening of conditions, results obtained using a sequential limit scheme usually correspond to that obtained under a joint limit theory (i.e., an asymptotic theory which is obtained by taking the limits of all indices jointly or simultaneously). While a general treatment of the conditions under which sequential and pathwise limit theories coincide for the IV problem studied here is beyond the scope of this paper; we note that, for the special case of an orthonormal SEM with Gaussian errors, it is easy to show that the sequential limit results given in Theorem 4.1 do in fact correspond with results obtained from a pathwise asymptotic procedure which takes T, k_{21}, and $\mu'\mu$ to infinity simultaneously, under the additional conditions that $\frac{\mu'\mu}{k_{21}}=\tau^2+Op\left(k_{21}^{-2}\right)$ and that $\frac{k}{T}=\alpha_T$, where $k=k_1+k_2$ and where α_T is a sequence of constants such that $0<\inf_T\alpha_T\leq\sup_T\alpha_T\leq 1$[6]. To see this, consider, for example, the special case where the

[6] The latter condition helps to ensure that, in the parameter sequence under consideration, the number of exogenous variables in the system never exceeds the sample size.

exogenous regressors are assumed to be fixed in repeated samples, and are orthonormalized, so that:

$$\overline{Z}'\overline{Z}=\begin{pmatrix} X'X & X'Z_1 & X'Z_2 \\ Z_1'X & Z_1'Z_1 & Z_1'Z_2 \\ Z_2'X & Z_2'Z_1 & Z_2'Z_2 \end{pmatrix} = \begin{pmatrix} TI_{k_1} & 0 & 0 \\ 0 & TI_{k_{21}} & 0 \\ 0 & 0 & TI_{k_{22}} \end{pmatrix}. \tag{25}$$

In addition, impose a Gaussian error assumption, so that $(u_t,\, v_t)' \equiv i.i.d.\ N(0, \Sigma)$. It is easy to see that if the local-to-zero assumption (Assumption 1) is maintained, then:

$$\begin{aligned}
\widehat{\beta}_{IV} - \beta_0 &= (y_2'(P_H - P_X)y_2)^{-1}(y_2'(P_H - P_X)u) \\
&= \left(y_2'M_X Z_1(Z_1'M_X Z_1)^{-1}Z_1'M_X y_2\right)^{-1}\left(y_2'M_X Z_1(Z_1'M_X Z_1)^{-1}Z_1'M_X u\right) \\
&= \left[\left(\left(T^{-1}Z_1'Z\right)C + T^{-\frac{1}{2}}Z_1'v\right)'\left(\left(T^{-1}Z_1'Z\right)C + T^{-\frac{1}{2}}Z_1'v\right)\right]^{-1} \times \\
&\quad \left[\left(\left(T^{-1}Z_1'Z\right)C + T^{-\frac{1}{2}}Z_1'v\right)'\left(T^{-\frac{1}{2}}Z_1'u\right)\right] \\
&\equiv \left(\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\right)\left[\left(C_1/\sigma_{vv}^{\frac{1}{2}} + Z_{v,1}\right)'\left(C_1/\sigma_{vv}^{\frac{1}{2}} + Z_{v,1}\right)\right]^{-1}\left[\left(C_1/\sigma_{vv}^{\frac{1}{2}} + Z_{v,1}\right)'Z_{u,1}\right] \\
&\equiv \left(\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\right)\left[(\mu + Z_{v,1})'(\mu + Z_{v,1})\right]^{-1}\left[(\mu + Z_{v,1})'Z_{u,1}\right],
\end{aligned} \tag{26}$$

where $\mu = C_1/\sigma_{vv}^{\frac{1}{2}}$, $Z_{v,1} = T^{-\frac{1}{2}}Z_1'v\sigma_{vv}^{-\frac{1}{2}}$, and $Z_{u,1} = T^{-\frac{1}{2}}Z_1'u\sigma_{uu}^{-\frac{1}{2}}$, and where it is obvious that

$$\begin{pmatrix} Z_{u,1} \\ Z_{v,1} \end{pmatrix} \sim N\left(0, \left(\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \otimes I_{k_{21}}\right)\right). \tag{27}$$

Note first that expression (26) gives the exact finite sample distribution of $\widehat{\beta}_{IV} - \beta_0$ under the assumptions of orthonormalized exogenous regressors and Gaussian errors. Note also that the finite sample distribution here depends on k_{21}, since μ, $Z_{v,1}$, and $Z_{u,1}$ are each $k_{21} \times 1$ vectors. Comparing expression (26) with the limiting distribution of the IV estimator given in Lemma A1 of Appendix A, we see that the finite sample distribution in this case has the same form as the asymptotic distribution, as has been pointed out by Staiger and Stock (1997). Moreover, note that the finite sample distribution of $\widehat{\beta}_{IV} - \beta_0$ in this case does not depend on the sample size T, since $\mu'\mu = C_1'C_1/\sigma_{vv}$ does not depend on T, and the joint distribution of $Z_{v,1}$ and $Z_{u,1}$ does not depend on T. As a consequence, for a given value of k_{21} and for all T, the formula for the exact bias function in this case is given by specializing expression (5) above to the orthonormal case. Similarly, the formula for the exact MSE function is given by specializing expression (6) to the orthonormal case. Hence, it follows trivially in this case that given the condition $\frac{\mu'\mu}{k_{21}} = \tau^2 + Op\left(k_{21}^{-2}\right)$, the asymptotic bias and MSE obtained by simultaneously taking T, k_{21}, and $\mu'\mu$ to infinity is precisely that which is obtained by a sequential limit procedure whereby T is first taken to infinity, and k_{21} and $\mu'\mu$ are subsequently taken to infinity.

(v) It is also of interest to compare and contrast our framework with the asymptotic framework employed by Morimune (1983), Bekker (1994), and Hahn (1997). These papers also examine properties of single-equation estimators when there is a large number of instruments, but none of these papers is concerned with the problem of weak instruments. More specifically, using the language of Phillips and Moon (1999), the asymptotic framework of Morimune (1983), Bekker (1994) and Hahn (1997) may be described as being based on a diagonal path limit scheme, whereby k_{21} increases monotonically as T increases such that $k_{21}/T \to \alpha$

for some real constant $\alpha \in (0,1)$. This is in contrast to our sequential limit scheme where the limit with respect to k_{21} is taken subsequent to the passage to infinity of the index T. One advantage of our sequential limit approach is that it allows us to make more general assumptions on both the exogenous regressors and the disturbances of the SEM, although this level of generality is does not need to come at the price of using a sequential limit approach, at leaset in the following sense. Under some stronger but more primitive conditions than those used in this paper, the bias-corrected estimators proposed here are also consistent under a pathwise asymptotic scheme whereby the number of instruments is taken to approach infinity as a function of the sample size, a result which is proven elsewhere (i.e. see Chao and Swanson (2001)).[7].

(vi) As can be seen from the proof of Theorem 4.1 and of Lemma A7 given in the Appendices, the approximate formulae for the ABIAS and AMSE in expressions (16) and (18) above are derived using a Laplace approximation of the confluent hypergeometric function which holds as $\mu'\mu$, $k_{21} \to \infty$ such that $\frac{\mu'\mu}{k_{21}} = \tau^2 + O\left(k_{21}^{-2}\right)$. This approximation is carried out by first rewriting the confluent hypergeometric function in its integral representation. To illustrate the main ideas of our procedure, we take the asymptotic expansion given in part (a) of Lemma A7 as an example and note that, in this case, the approximation is extracted by applying Laplace's method to the representation

$$\begin{aligned}
& {}_1F_1\left(\frac{k_{21}}{2}-1;\ \frac{k_{21}}{2};\ \frac{\mu'\mu(\tau^2,k_{21})}{2}\right)\exp\left\{-\frac{\mu'\mu(\tau^2,k_{21})}{2}\right\} \\
= \ & \frac{\Gamma\left(\frac{k_{21}}{2}\right)}{\Gamma\left(\frac{k_{21}}{2}-1\right)}\int_0^1 \exp\left\{\frac{\mu'\mu(\tau^2,k_{21})}{2}(t-1)\right\} t^{(k_{21}-4)/2}dt. \qquad (28)
\end{aligned}$$

Loosely speaking, this is done by expanding the integrand of the integral expression on the right-hand side of equation (28) above in a Taylor series about the maximum of the integrand and integrating this series term-by-term. Note, however, that the Laplace approximation used here differs from that which is commonly applied in Bayesian analysis to compute posterior means or marginal posterior densities. (See Kass, Tierney, and Kadane (1990) and Chao and Phillips (1998) for examples of the application of the Laplace approximation in Bayesian analysis.) More specifically, whereas in the more common applications of Laplace's method, the maximum of the integrand typically occurs in the interior of the domain of integration, the maximum of the integrand here occurs at the upper boundary, i.e., at $t = 1$. Hence, the term involving the first derivative of the integrand in its Taylor expansion does not vanish here as it does in the case of an interior maximum. (See de Bruijn, 1961, for a more detailed discussion of the Laplace approximation in the case where the maximum lies on the boundary of the domain of integration.)

Note also that while, for the purpose of this paper, the asymptotic expansions of the confluent hypergeometric functions given in Lemma A7 are only intermediate results used to obtain approximate formulae for the ABIAS and AMSE functions of the IV estimator; these expansions may be of interest in their own right in other contexts. In particular, these results add to the already extensive literature in applied mathematics on the asymptotic behavior of the confluent hypergeometric function and, in fact, extend the particular result given by equation (4.3.8) of Slater (1960) in at least two ways. First, a careful reading of the steps leading up to equation (4.3.8) of Slater (1960) show that it 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

[7]The pathwise asymptotic results are available from the authors upon request.

based on the Laplace approximation, do apply in this case (as is needed if these results are going to provide useful approximations for the ABIAS and AMSE functions). Secondly, (translated into our notations) the derivation of equation (4.3.8) in Slater (1960) imposes the exact relation $\frac{\mu'\mu}{k_{21}} = \tau^2$, so that $\mu'\mu$ grows to infinity as an exact linear function of k_{21}. On the other hand, our results assume the less restrictive condition which does not require $\mu'\mu$ to be an exact linear function of k_{21} but rather allow the ratio, $\frac{\mu'\mu}{k_{21}}$, to deviate from τ^2 by a remainder term $R(k_{21}) \equiv \frac{\mu'\mu}{k_{21}} - \tau^2$, which, in turn, vanishes at the rate k_{21}^{-2} as $k_{21} \to \infty$.
(vii) While Assumption 4 requires that $\tau^2 > 0$, it is easy to see, by following the proof of Theorem 4.1, that the bias and MSE expansion given by expressions (16) and (18) are valid even for $\tau^2 = 0$. However, the condition $\tau^2 > 0$ is assumed because, as explained in Remark 5.8 (i) of the next section, τ^2 must not be zero if we are going to be able to construct consistent estimators of the lead term of the bias and MSE expansions and to construct bias-adjusted estimators which estimate consistently the structural coefficient β. See Remark 5.8 (i) for further discussion.
(viii) Note, of course, that Assumption 4 is made so we can carry out a second order approximation of the ABIAS and AMSE functions, with an approximation error of order $O\left(k_{21}^{-2}\right)$. It is easily seen that Assumption 4, in fact, gives the slowest possible rate of decay of the remainder term $R(k_{21})$ under which an approximation error of order $O\left(k_{21}^{-2}\right)$ can be achieved. Hence, suppose instead that we assume $\frac{\mu'\mu}{k_{21}} = \tau^2 + O\left(k_{21}^{-(1+\delta)}\right)$ where $0 < \delta < 1$; then, it can be shown easily that a second order approximation will now have approximation error of order $O\left(k_{21}^{-(1+\delta)}\right)$ and that the approximation error in this case will be dominated by a term involving the remainder component $R(k_{21})$.
(ix) In the literature on finite sample distributions of IV estimators, it is customary to write the formulae for the bias and MSE as functions of $\mu'\mu$ and k_{21} (e.g. see Richardson and Wu (1971)). We adopt this convention in Section 3. However, the approximations given in Theorem 4.1 suggest that it may also be sensible to instead think of the bias and MSE as functions of k_{21} and the ratio $\frac{\mu'\mu}{k_{21}}$ (particularly since our numerical results in Section 6 suggest that expansions where both $\mu'\mu$ and k_{21} are allowed to grow to infinity in the way stipulated by Assumption 4 actually results in a better approximation than an alternative expansion which only allows k_{21} to grow to infinity while keeping $\mu'\mu$ fixed). Hence, it is of interest to examine the bias and MSE functions when they are written in terms of k_{21} and the ratio $\frac{\mu'\mu}{k_{21}}$. To proceed, set $\tau^2 = \frac{\mu'\mu}{k_{21}}$ and reparameterize the asymptotic bias and MSE given in Theorems 3.1 and 3.2 in terms of τ^2 and k_{21} as:

$$
\begin{aligned}
b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho e^{-\frac{\tau^2 k_{21}}{2}}\ {}_1F_1\left(\frac{k_{21}}{2}-1; \frac{k_{21}}{2}; \frac{\tau^2 k_{21}}{2}\right), \qquad (29)\\
m_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\tau^2 k_{21}}{2}}\left[\frac{1}{\rho^2}\left(\frac{1}{k_{21}-2}\right)\ {}_1F_1\left(\frac{k_{21}}{2}-1; \frac{k_{21}}{2}; \frac{\tau^2 k_{21}}{2}\right)\right.\\
&\quad \left. + \left(\frac{k_{21}-3}{k_{21}-2}\right)\ {}_1F_1\left(\frac{k_{21}}{2}-2; \frac{k_{21}}{2}; \frac{\tau^2 k_{21}}{2}\right)\right] \qquad (30)
\end{aligned}
$$

Based on these formulae, we obtain the following results.

Theorem 4.4: *Let $b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ and $m_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ be as defined in expressions (29) and (30), respectively. Then, it follows that for $\sigma_{uv} \neq 0$:*

(a) *$|b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})|$ is a monotonically decreasing function of τ^2 for fixed k_{21}.*
(b) *$|b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})|$ is a monotonically increasing function of k_{21} for fixed τ^2.*

(c) $m_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ *is a monotonically decreasing function of* τ^2 *for fixed* k_{21}.

The results of Theorem 4.4 provide some characterization of the partial effects that changes in τ^2 and k_{21} have on the absolute magnitude of the ABIAS and AMSE. Note, in particular, that parts (a) and (c) of the theorem imply that, for k_{21} fixed, an increase in τ^2 decreases the ABIAS and AMSE of the IV estimator. This result is, of course, intuitive since $\tau^2 = \frac{\mu'\mu}{k_{21}}$ may be loosely interpreted as the population analogue of the first stage F-statistic so that a higher value of τ^2 implies a more relevant set of instruments. In addition, note that part (b) of Theorem 4.4 implies that, given τ^2, an increase in k_{21} increases the ABIAS. No simple statement can be made, however, with respect to the partial effect that an increase in k_{21} has on the AMSE for a fixed value of τ^2, as the effect can be either positive or negative depending on the values of the parameters σ_{uu}, σ_{uv}, σ_{vv}, and τ^2. This is because, given τ^2, an increase in k_{21} increases the ABIAS but decreases the (asymptotic) variance of the IV estimator, so that the direction of the net effect is in general ambiguous.

5 Estimation of Bias and MSE and Bias Correction

5.1 Consistent Estimation of the Bias and MSE

In this subsection, we obtain consistent estimators for the lead terms of the ABIAS and AMSE. 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 obtain consistent estimation of the quantities σ_{vv}, σ_{uv}, σ_{uu}, and $(1 + \tau^2)$ under the sequential limit approach.

Lemma 5.1: *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:*

(a) $\widehat{\sigma}_{vv,i} \xrightarrow{p} \sigma_{vv}$, *for* $i = 1, 2$;

(b) $W_{k_{21},T} \xrightarrow{p} 1 + \tau^2$;

(c) $\widehat{\sigma}_{uv,i} \xrightarrow{p} \sigma_{uv}$, *for* $i = 1, 2$;

(d) $\widehat{\sigma}_{uu,i} \xrightarrow{p} \sigma_{uu}$, *for* $i = 1, 2$.

Based on these estimators, we propose four different estimators for the ABIAS and six different estimators for the AMSE, as follows:

$$\widehat{BIAS}_i = \frac{\widehat{\sigma}_{uv,i}}{\widehat{\sigma}_{vv,i}}\left(\frac{1}{W_{k_{21},T}}\right), \text{ for } i = 1, 2; \tag{31}$$

$$\begin{aligned} \widetilde{BIAS}_i &= \frac{\widehat{\sigma}_{uv,i}}{\widehat{\sigma}_{vv,i}}\left[\left(\frac{1}{W_{k_{21},T}}\right) - \frac{1}{k_{21}}\left(\frac{1}{W_{k_{21},T}}\right)\left\{2 - 4\left(\frac{1}{W_{k_{21},T}}\right) + 2\left(\frac{1}{W_{k_{21},T}}\right)^2\right\}\right] &(32)\\ &= \frac{\widehat{\sigma}_{uv,i}}{\widehat{\sigma}_{vv,i}}\left[\left(\frac{1}{W_{k_{21},T}}\right) - \frac{2}{k_{21}}\left(\frac{1}{W_{k_{21},T}}\right)\left(\frac{W_{k_{21},T} - 1}{W_{k_{21},T}}\right)^2\right], \text{ for } i = 1, 2; &(33)\end{aligned}$$

$$\widehat{MSE}_i = \frac{\widehat{\sigma}_{uv,i}^2}{\widehat{\sigma}_{vv,i}^2}\left(\frac{1}{W_{k_{21},T}}\right)^2, \text{ for } i = 1, 2; \tag{34}$$

$$\widetilde{MSE}_i = \frac{\widehat{\sigma}^2_{uv,i}}{\widehat{\sigma}^2_{vv,i}} \left[\left(\frac{1}{W_{k_{21},T}} \right)^2 + \frac{1}{k_{21}} \left(\frac{\widehat{\sigma}_{uu,i}\widehat{\sigma}_{vv,i} - \widehat{\sigma}^2_{uv,i}}{\widehat{\sigma}^2_{uv,i}} \right) \left(\frac{1}{W_{k_{21},T}} \right) + \frac{1}{k_{21}} \left(\frac{1}{W_{k_{21},T}} \right) \left(1 - \frac{7}{W_{k_{21},T}} + \frac{12}{W^2_{k_{21},T}} - \frac{6}{W^3_{k_{21},T}} \right) \right], \text{ for } i = 1, 2; \quad (35)$$

$$\overline{MSE}_i = \frac{\widehat{\sigma}^2_{uv,i}}{\widehat{\sigma}^2_{vv,i}} \left[\left(\frac{1}{W_{k_{21},T}} \right)^2 + \frac{1}{k_{21}} \left(\frac{\widehat{g}_{11}\widehat{g}_{22} - \widehat{g}^2_{12}}{\widehat{\sigma}^2_{uv,i}} \right) \left(\frac{1}{W_{k_{21},T}} \right) + \frac{1}{k_{21}} \left(\frac{1}{W_{k_{21},T}} \right) \left(1 - \frac{7}{W_{k_{21},T}} + \frac{12}{W^2_{k_{21},T}} - \frac{6}{W^3_{k_{21},T}} \right) \right], \text{ for } i = 1, 2. \quad (36)$$

Note that the difference between the "hat" estimators and the "tilde" estimators is that the "hat" estimators are constructed based only on the lead term of the expansions given in Theorem 4.1 while the "tilde" estimators make use of both the lead term and the second order term. In addition, the difference between $\widetilde{MSE}_i$ and $\overline{MSE}_i$ lies in the fact that, given the equivalence $\sigma_{uu}\sigma_{vv} - \sigma^2_{uv} = g_{11}g_{22} - g^2_{12}$, there are two ways we can estimate the quantity $\sigma_{uu}\sigma_{vv} - \sigma^2_{uv}$ (i.e. we can either estimate $\sigma_{uu}\sigma_{vv} - \sigma^2_{uv}$ directly or estimate it indirectly as $g_{11}g_{22} - g^2_{12}$). The next theorem derives the probability limits of the estimators given by expressions (31)-(36).

Theorem 5.2: *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:*

(a) $\widehat{BIAS}_i \xrightarrow{p} \sigma^{1/2}_{uu}\sigma^{-1/2}_{vv}\rho\left(\frac{1}{1+\tau^2}\right)$, *for* $i = 1, 2$;

(b) $\widetilde{BIAS}_i \xrightarrow{p} \sigma^{1/2}_{uu}\sigma^{-1/2}_{vv}\rho\left(\frac{1}{1+\tau^2}\right)$, *for* $i = 1, 2$;

(c) $\widehat{MSE}_i \xrightarrow{p} \sigma_{uu}\sigma^{-1}_{vv}\rho^2\left(\frac{1}{1+\tau^2}\right)^2$, *for* $i = 1, 2$;

(d) $\widetilde{MSE}_i \xrightarrow{p} \sigma_{uu}\sigma^{-1}_{vv}\rho^2\left(\frac{1}{1+\tau^2}\right)^2$, *for* $i = 1, 2$;

(e) $\overline{MSE}_i \xrightarrow{p} \sigma_{uu}\sigma^{-1}_{vv}\rho^2\left(\frac{1}{1+\tau^2}\right)^2$, *for* $i = 1, 2$.

Remark 5.3: (i) The estimators defined in equations (31)-(36) are all weakly consistent, in the sense that each bias estimator converges in probability to the lead term of the bias expansion given in (16), while each MSE estimator converges in probability to the lead term of the MSE expansion given by (18). These results suggest that there is information which can be exploited when a large number of weakly correlated instruments are available.

(ii) It is of interest to analyze the asymptotic properties of our bias and MSE estimators in the conventional framework, where the usual identification condition is assumed to hold. Hence, in place of the local-to-zero condition of Assumption 1**,** we make the alternative identification assumption:

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

In order to obtain the probability limits of our estimators under Assumption 1*, we first give a Lemma which contains limiting results for our estimators of the parameters σ_{vv}, σ_{uv}, σ_{uu}, and for the Wald statistic for testing instrument relevance, $W_{k_{21},T}$.

Lemma 5.4: *Suppose that Assumptions 1* and 2 hold. Then, as $T \to \infty$, the following limit results hold:*

(a) $\widehat{\sigma}_{vv,1} \xrightarrow{p} \sigma_{vv}$;

(b) $\widehat{\sigma}_{vv,2} \xrightarrow{p} \Pi'\Omega\Pi + \sigma_{vv}$;

(c) $W_{k_{21},T} = Op(T)$;
(d) $\widehat{\sigma}_{uv,i} \xrightarrow{p} \sigma_{uv}$, *for* $i = 1, 2$;
(e) $\widehat{\sigma}_{uu,i} \xrightarrow{p} \sigma_{uu}$, *for* $i = 1, 2$.

In view of expressions (31)-(36), the next theorem follows as an immediate consequence of Lemma 5.4 and the Slutsky's theorem:

Theorem 5.5: *Suppose that Assumptions 1* and 2 hold. Then, as* $T \to \infty$, *the following limit results hold,*
(a) $\widehat{BIAS}_i \xrightarrow{p} 0$, *for* $i = 1, 2$;
(b) $\widetilde{BIAS}_i \xrightarrow{p} 0$, *for* $i = 1, 2$;
(c) $\widehat{MSE}_i \xrightarrow{p} 0$, *for* $i = 1, 2$;
(d) $\widetilde{MSE}_i \xrightarrow{p} 0$, *for* $i = 1, 2$;
(e) $\overline{MSE}_i \xrightarrow{p} 0$, *for* $i = 1, 2$.

Note that, all of these bias and MSE estimators approaches zero, as $T \to \infty$ (as they should in the case of full identification since the IV estimator is weakly consistent in this case). These results suggest that our estimators behave in a reasonable manner even in the conventional case where instruments are fully relevant (i.e. when Assumption 1* holds).
(iii) One possible use of the MSE estimators given above is as descriptive statistics for assessing instrument relevance. For such an application, it might be convenient to construct measures of the AMSE of the IV estimator relative to that of the OLS estimator, viz

$$\widehat{RM}_i = \left(\frac{1}{W_{k_{21},T}}\right)^2, \text{ for } i = 1, 2, \text{ and} \tag{37}$$

$$\begin{aligned} \widetilde{RM}_i \quad = \quad & \left[\left(\frac{1}{W_{k_{21},T}}\right)^2 + \frac{1}{k_{21}}\left(\frac{\widehat{\sigma}_{uu,i}\widehat{\sigma}_{vv,i} - \widehat{\sigma}^2_{uv,i}}{\widehat{\sigma}^2_{uv,i}}\right)\left(\frac{1}{W_{k_{21},T}}\right)\right. \\ & \left. + \frac{1}{k_{21}}\left(\frac{1}{W_{k_{21},T}}\right)\left(1 - \frac{7}{W_{k_{21},T}} + \frac{12}{W^2_{k_{21},T}} - \frac{6}{W^3_{k_{21},T}}\right)\right], \text{ for } i = 1, 2, \end{aligned} \tag{38}$$

where $\widehat{RM}_i$ and $\widetilde{RM}_i$ $(i = 1, 2)$ are obtained from $\widehat{MSE}_i$ and $\widetilde{MSE}_i$ $(i = 1, 2)$, respectively, by dividing by $\widehat{\sigma}^2_{uv,i}/\widehat{\sigma}^2_{vv,i}$ $(i = 1, 2)$, and where $\widehat{\sigma}^2_{uv,i}/\widehat{\sigma}^2_{vv,i}$ $(i = 1, 2)$ are consistent estimators of $\sigma^2_{uv}/\sigma^2_{vv} = \sigma_{uu}\sigma_{vv}\rho^2$ (which is the AMSE of the OLS estimator under the local-to-zero assumption). Note, in particular, that if the set of instruments are fully relevant then one would expect to observe values of these relative MSE measures close to zero. On the other hand, if the instruments are weak, then the statistics should take values closer to, and possibly in excess of, unity. In future research, we plan to explore more fully the use of the bias and MSE measures given above, both for the *ex-post* assessment of instrument relevance and possibly also for addressing the important problem of instrument selection.[8]

[8] In a simulating paper, Hall, Rudebusch, and Wilcox (1996) caution against the use of first-stage F-statistic as a screening device for instrument relevance. They point out that a high value of the first-stage F-statistic might not always indicate that the (population) correlation between the instrument Z_t and the endogenous regressor y_{2t} is high. Rather, a high F-statistic might instead identify a spuriously high sample correlation between Z_t and the disturbance of the structural equation u_t; which could occur because, given the endogeneity of y_{2t}, a high sample correlation between y_{2t} and Z_t, as indicated by the F-statistic,

5.2 Bias Correction

The results of the last subsection is that they can be used to construct bias-adjusted OLS and IV estimators, which give consistent estimation of the structural coefficient β within our sequential limit framework.[9] In particular, we propose five alternative bias-corrected estimators:

$$\widetilde{\beta}_{OLS,i} = \widehat{\beta}_{OLS} - \frac{\widehat{\sigma}_{uv,i}}{\widehat{\sigma}_{vv,i}}, \text{ for } i = 1, 2; \tag{39}$$

$$\widetilde{\beta}_{IV} = \widehat{\beta}_{IV} - \widehat{BIAS}_1; \text{ and} \tag{40}$$

$$\widetilde{\widetilde{\beta}}_{IV,i} = \widehat{\beta}_{IV} - \widetilde{BIAS}_i, \text{ for } i = 1, 2. \tag{41}$$

In Section 7, we report the results of a Monte Carlo study which evaluates the performance of these bias-corrected estimators vis-a-vis the uncorrected OLS and IV estimators. First, however, we present two theorems which give, respectively, the probability limits of the bias-corrected estimators under the local-to-zero condition of Assumption 1 and under the more conventional full-identification condition given by Assumption 1^*.

Theorem 5.6: *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:*

(a) $\widetilde{\beta}_{OLS,i} \xrightarrow{p} \beta_0$ *for* $i = 1, 2$;

(b) $\widetilde{\beta}_{IV} \xrightarrow{p} \beta_0$; and

(c) $\widetilde{\widetilde{\beta}}_{IV,i} \xrightarrow{p} \beta_0$ *for* $i = 1, 2$.

Theorem 5.7: *Suppose that Assumptions 1* and 2 hold. Then, as $T \to \infty$, the following limit results hold:*

(a) $\widetilde{\beta}_{OLS,1} \xrightarrow{p} \beta_0 - \frac{\sigma_{uv}}{\sigma_{vv}} \left(\frac{\Pi'\Omega\Pi}{\Pi'\Omega\Pi + \sigma_{vv}} \right)$;

(b) $\widetilde{\beta}_{OLS,2} \xrightarrow{p} \beta_0$;

(c) $\widetilde{\beta}_{IV} \xrightarrow{p} \beta_0$;

(d) $\widetilde{\widetilde{\beta}}_{IV,i} \xrightarrow{p} \beta_0$ *for* $i = 1, 2$.

Remark 5.8: (i) Note that although the bias-corrected estimators given above are consistent in our framework, uncorrected OLS and IV estimators are not. It should also be emphasized that if $\tau^2 = 0$, then the bias-adjusted estimators introduced above would not consistently estimate the structural coefficient β. This is due to the fact that the bias-adjusted estimators make use of covariance estimators of the form: $s_{uv,i} = \frac{(y_1 - y_2\widehat{\beta}_{IV})' M_i y_2}{T}$, for $i = 1$ or 2; and, $s_{uv,i} \xrightarrow{p} \sigma_{uv} \left(\frac{\tau^2}{1+\tau^2} \right)$ for $i = 1, 2$ (as shown in the proof of part (c) of Lemma 5.1). Consequently, if $\tau^2 = 0$, $s_{uv,i} \xrightarrow{p} 0$ for $i = 1, 2$; hence, asymptotically neither $s_{uv,1}$

might give rise to a spurious correlation between Z_t and u_t. Under the latter scenario, using the first-stage F-statistic as a screening device might lead to severe pre-test bias, as documented in their Monte Carlo study. In such a case, it may be fruitful to use instead the formulae for the (absolute) bias and MSE given in expressions (31)-(36) as diagnostic statistics for assessing instrument relevance since a measure of the covariance between y_{2t} and u_t (either $\widehat{\sigma}_{uv,1}$ or $\widehat{\sigma}_{uv,2}$) appear in each of our estimators for the (absolute) bias and MSE. Since a high sample correlation between Z_t and u_t is more likely to occur if the degree of endogeneity of y_{2t} is high, our bias and MSE formulae implicitly adjust for the possible distortions caused by a spurious (in sample) correlation between Z_t and u_t whereas the first-stage F statistic does not.

Other interesting measures of instrument relevance are discussed in Shea (1997) and Hall ... .

[9] Other recent work in the area of bias correction is contained in Hausman, Hahn, and Kuersteiner (2001), and the references cited therein, for example.

nor $s_{uv,2}$ carries any information about the value of σ_{uv} and, thus, cannot be adjusted to obtain consistent estimators. Indeed, since $\tau^2 = 0$ arises either because the model is unidentified in the traditional sense (i.e., $C = 0$) or, more generally, because all but a finite number of instruments are completely uncorrelated with the endogenous explanatory variable as the number of instruments approaches infinity, we would not expect consistent estimation of β to be possible when $\tau^2 = 0$. Our results suggest that if one is faced with a situation where only a great many weak instruments are available; then, it may still be worthwhile to make use of these poor quality instruments in constructing bias- corrected estimators of β, so long as the instruments are not completely uncorrelated with the endogenous explanatory variable (i.e., so long as the situation is not well-modeled by the case $\tau^2 = 0$).
(ii) Theorem 5.7 shows that in the conventional case where the instruments are fully relevant, all but one of the bias-corrected estimators are still consistent. Indeed, only $\widetilde{\beta}_{OLS,1}$ is inconsistent under Assumption 1*, and the reason for its inconsistency is that in this case it can easily be shown that $\widehat{\beta}_{OLS} \xrightarrow{p} \beta_0 + \frac{\sigma_{uv}}{\Pi'\Omega\Pi+\sigma_{vv}}$, whereas the bias-correction factor $-(\widehat{\sigma}_{uv,1}/\widehat{\sigma}_{vv,1})$ is a consistent estimator which converges in probability to $-\sigma_{uv}/\sigma_{vv}$, so that it does not cancel out the bias of the uncorrected OLS estimator, $\widehat{\beta}_{OLS}$.

6 Numerical Results

In this section, we present some numerical results on the ABIAS and AMSE. Our numerical calculations are based on the canonical SEM; that is, an SEM as described in Section 2 above, except that the reduced form error covariance matrix is taken to be an identity matrix (i.e., $G = I$ using the notation introduced in Remark 3.3(vi)). Note that the canonical model is often used in the finite sample literature on single-equation estimators to obtain both analytical and numerical/Monte Carlo results. (See, for example, Richardson and Wu (1971), Mariano and McDonald (1979), and Phillips (1983, 1984, 1985).) As explained in Phillips (1983), the non-canonical SEM as described in Section 2 can always be transformed into a canonical model by applying the appropriate standardizing transformations, so that the canonical model provides a representation which preserves many of the important features of the original (untransformed) SEM but at the same time reduces the parameter space of that model to an essential set. As a consequence, the canonical model is ideally suited for the design of numerical and Monte Carlo experiments.

We wish first to present some graphical illustrations of the shape of the ABIAS and AMSE functions, reparameterized in terms of τ^2 and k_{21}, (i.e., the ABIAS and AMSE functions given by expressions (29) and (30)). To proceed, note first that for a canonical model $\sigma_{vv} = 1$, $\sigma_{uu} = 1 + \beta^2$, and $\rho = -\beta/\sqrt{1+\beta^2}$, so that in this case the elements of the error covariance matrix of the structural model (i.e., Σ) depend only on the value of the structural coefficient β. It follows that for a canonical model the bias and MSE functions given by expressions (29) and (30) depend only on the quantities β, τ^2, and k_{21}. To get a sense of how the bias and MSE vary as a function of $s = \left(1+\tau^2\right)^{-1}$ and k_{21}, we set $\beta = -0.5$ and plot, in Figure 1 (in the back of the paper), the relative bias $b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})/b_{\widehat{\beta}_{OLS}}$ and the relative MSE $m_{\widehat{\beta}_{IV}}(\tau^2, k_{21})/m_{\widehat{\beta}_{OLS}}$ against k_{21} for s fixed and also plot, in Figure 2, $b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})/b_{\widehat{\beta}_{OLS}}$ and $m_{\widehat{\beta}_{IV}}(\tau^2, k_{21})/m_{\widehat{\beta}_{OLS}}$ against s for k_{21} fixed[10]. Look

[10]Note that the values of the functions $b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ and $m_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$, plotted in these graphs, are computed on the basis of the *actual* (as opposed to the approximate) formulae given in expressions (29) and (30). Moreover, $b_{\widehat{\beta}_{OLS}}$ and $m_{\widehat{\beta}_{OLS}}$ here denote the asymptotic bias and MSE of the OLS estimator under Assumption 1, so that for the canonical model

first at the top diagram in Figure 1, we see that for a fixed value of s (or, alternatively, a fixed value of τ^2) the value of the relative bias increases rapidly at first but levels off quickly toward the value $s = \left(1+\tau^2\right)^{-1}$, which is, in fact, the value of the lead term of our approximation of the relative bias based on Assumption 4. Although, in this first diagram, we have only plotted the relative bias curve for the cases $s = 1/2$, $1/4$, and $1/8$ (corresponding to $\tau^2 = 1$, 3, and 7), the picture which emerges is actually quite representative of that for other values of s (or τ^2) as well. Indeed, for the boundary case where $s = 1$ (or $\tau^2 = 0$), it is easy to see that the convergence to the first order term becomes in some sense "instantaneous" since the relative bias function in this case is unity for all values of k_{21}. Now, looking at the two bottom diagrams of Figure 1, we see that for a fixed value of s (or τ^2), the values of the MSE function also converge quickly, as k_{21} increases, but to the value $s^2 = \left(1+\tau^2\right)^{-2}$, which in turn is the value of the lead term of the MSE expansion (18), when the latter is standardized by $m_{\widehat{\beta}_{OLS}}$[11]. To illustrate with a specific example, we note that for $s^2 = \left(1+\tau^2\right)^{-2} = 1/4$, the relative MSE is 1.1484 for $k_{21} = 4$, is 0.8512 for $k_{21} = 5$, and is 0.2692 for $k_{21} = 100$. Overall, the fairly rapid convergence of both the relative bias and the relative MSE to the values of the lead terms in our expansions suggest that, even for moderate k_{21}, a first order approximation based on the results of Theorem 4.1 is likely to be quite satisfactory.

Next, we turn our attention to Figure 2, where, for fixed values of k_{21}, relative bias is plotted against $s = \left(1+\tau^2\right)^{-1}$ in the top diagram while relative MSE is plotted against $s = \left(1+\tau^2\right)^{-1}$ in the two bottom diagrams. We see in Figure 2 that for a fixed value of k_{21}, both the relative bias and the relative MSE increase as s increases (or as τ^2 decreases). This behavior is to be expected since, loosely speaking, τ^2 is the population analogue of the F-statistic for testing instrument relevance; hence, the smaller the value of τ^2, the weaker are the instruments. In addition, note that the observation in Figure 2 that the relative bias and MSE are increasing functions of s for k_{21} fixed is also consistent with the depiction in Figure 1, where the relative bias and MSE curves are shown to shift upward as s increases. Furthermore, the top diagram of Figure 2 show that the entire relative bias curve (as a function of s) shifts upward if we increase the value of k_{21}. Again, this is consistent with the top diagram in Figure 1, which shows relative bias to be an increasing function of k_{21}.

We now turn to an evaluation of the accuracy of the approximations given in Section 4. Note that it may be of interest to compare the accuracy of this type of approximation (henceforth referred to as Approximation Scheme 1) with an alternative approximation which is obtained by taking only k_{21} to infinity. Under this latter asymptotic scheme (which we shall refer to as Approximation Scheme 2); we obtain, up to a remainder term of order $O\left(k_{21}^{-2}\right)$, the following expansions of the formulae for the bias and MSE:

$b_{\widehat{\beta}_{OLS}} = \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho = -\beta$ and $m_{\widehat{\beta}_{OLS}} = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 = \beta^2$.

[11] A few words of caveat must be added with respect to Figure 1. While the two bottom diagrams in Figure 1 depict the graphs of the relative MSE in cases where it is a monotonically decreasing function of k_{21} for s (or τ^2) fixed, it should be noted that, unlike the relative bias curve (which is always monotonically increasing with respect to k_{21} except in the boundary case where $s = 1$, or $\tau^2 = 0$), the shape of the relative MSE curve depends, in general, on values of the parameters β and s (or τ^2) and may not be everywhere decreasing with respect to k_{21}. In fact, it is not hard to see that given some values of k_{21} and s (or τ^2), the relative MSE may be an increasing function of k_{21} in some regions. Note, however, that in spite of the fact that the relative MSE curve may have several different shapes, it does in all cases converge rapidly to the value of the lead term of our MSE expansion as k_{21} increases. As it is not feasible for us to give a whole menu of graphs depicting each possible shape of the relative MSE curve, the graphs depicted in the two bottom diagrams of Figure 1 should be interpreted as only giving only a flavor of the possible behavior of the relative MSE function.

$$b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1-\left(\frac{\mu'\mu}{k_{21}}\right)\right] + O(k_{21}^{-2}), \quad \text{and} \tag{42}$$

$$m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 + \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left[\left(\frac{1-\rho^2}{\rho^2}\right)\left(\frac{1}{k_{21}}\right) - 2\left(\frac{\mu'\mu}{k_{21}}\right)\right] + O(k_{21}^{-2}). \tag{43}$$

Given these formulae, our assessment and comparison of the two approximation schemes are based on variants of the following regressions:

Regression for Approximation Scheme 1

$$\begin{aligned} b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) &= \phi_0 + \phi_1\left[-\beta\left(1+\mu'\mu/k_{21}\right)^{-1}\right] + \phi_2\left[-\beta k_{21}^{-1}\left(1+\mu'\mu/k_{21}\right)^{-1}\right] + \\ &\quad \phi_3\left[-\beta k_{21}^{-1}\left(1+\mu'\mu/k_{21}\right)^{-2}\right] + \phi_3\left[-\beta k_{21}^{-1}\left(1+\mu'\mu/k_{21}\right)^{-3}\right] + error, \end{aligned} \tag{44}$$

$$\begin{aligned} m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) &= \pi_0 + \pi_1\left[\beta^2\left(1+\mu'\mu/k_{21}\right)^{-2}\right] + \pi_2\left[\left(1+\beta^2\right)k_{21}^{-1}\left(1+\mu'\mu/k_{21}\right)^{-1}\right] \\ &\quad +\pi_3\left[\beta^2 k_{21}^{-1}\left(1+\mu'\mu/k_{21}\right)^{-2}\right] + \pi_4\left[\beta^2 k_{21}^{-1}\left(1+\mu'\mu/k_{21}\right)^{-3}\right] \\ &\quad +\pi_5\left[\beta^2 k_{21}^{-1}\left(1+\mu'\mu/k_{21}\right)^{-4}\right] + error, \end{aligned} \tag{45}$$

Regression for Approximation Scheme 2

$$b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \phi_0^* + \phi_1^*\left[-\beta\right] + \phi_2^*\left[-\beta\mu'\mu/k_{21}\right] + error, \tag{46}$$

$$m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \pi_0^* + \pi_1^*\left[\beta^2\right] + \pi_2^*\left[1/k_{21}\right] + \pi_3^*\left[\beta^2\mu'\mu/k_{21}\right] + error. \tag{47}$$

In performing our regression analysis, we calculate values of the dependent variable for the equations above using the analytical formulae for the bias and MSE given by expressions (7) and (8) but specialized to the case of a canonical model. Moreover, note that the independent variables in the regressions above (i.e., the terms in square brackets) are simply the terms in our approximation of the bias and MSE given in Theorem 4.1, again specialized to the canonical case. Values for both the dependent and the independent variables are calculated for various values of β, $\mu'\mu$, and k_{21}. More specifically, the grid of values used is: $\beta = \{-0.5, -1.0, -1.5, ..., -10\}$, $\mu'\mu = \{0, 2, 4, 6, 8, ..., 100\}$, and $k_{21} = \{3, 5, 7, 9, 11, ..., 101\}$, so that in total 51000 observations are generated by taking all possible combinations[12].

Regression results are summarized in Table 1 (bias) and Table 2 (MSE), given in the back of the paper. Comparing first the numbers reported in column 1 of Table 1, which gives results for regression (44) above, with the numbers reported in column 3 of the same table, which gives results for regression (46); we see that the regression based on Approximation Scheme 1 (i.e., regression (44)) clearly fit the variations in the bias function much better than the regression based on Approximation Scheme 2 (i.e., regression (46)). In

[12]Note that, for our regressions analysis, we have chosen only negative values of β. This is because the MSE only depends on β^2, so that the sign of β does not matter. On the other hand, the bias function is perfectly symmetrical with respect to positive and negative values of β, so that the inclusion of positive values of β will only change the sign but will not alter any important qualitative aspect of our analysis.

particular, note that the adjusted R^2 (henceforth, $\overline{R}^2$) for the regression results reported in column 1 of Table 1 is 1.0000 while $\overline{R}^2$ for the regression results reported in column 3 is only 0.7165. Similar results are also obtained for the MSE, as can be seen by comparing columns 1 and 3 of Table 2, which report results from running the regressions described by expressions (45) and (47), respectively. Indeed, for the MSE, $\overline{R}^2$ is 0.9998 for the regression based on Approximation Scheme 1 but only 0.5494 for that based on Approximation Scheme 2. Moreover, comparable results are obtained if we run regressions based on the relative bias $b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})/b_{\widehat{\beta}_{OLS}}$ and the relative MSE $m_{\widehat{\beta}_{IV}}(\tau^2, k_{21})/m_{\widehat{\beta}_{OLS}}$, instead of the (absolute) bias $b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ and the (absolute) MSE $m_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$[13]. These results are given in columns 4 and 6 of Table 1 for relative bias and in columns 4 and 6 of Table 2 for relative MSE. Note again that, in both tables, the fit of the regression reported in column 4 is much better than that of the regression reported in column 6, so that Approximation Scheme 1 outperforms Approximation Scheme 2. Overall, our numerical results suggest that simply taking k_{21} to infinity (as in Approximation Scheme 2) will not allow one to obtain a reasonable approximation for the (asymptotic) bias and MSE of the IV estimator. Rather, much better approximations can be obtained via Approximation Scheme 1; that is, by expanding the bias and MSE functions in the manner prescribed by Assumption 4. Indeed, while numerical calculations given in the finite sample literature on single-equation estimators have tend to parameterize the bias and MSE functions in terms of k_{21} and $\mu'\mu$ (see, for example, Richardson and Wu, 1971), our results here suggest that both the bias and the MSE might also be viewed as functions of k_{21} and the ratio $\tau^2 = \frac{\mu'\mu}{k_{21}}$ rather than as functions of k_{21} and $\mu'\mu$. This is a point also made by Staiger and Stock (1997).

Returning to columns 1 of Tables 1 and 2 and comparing regression results reported there with the theoretical results obtained in Theorem 4.1, we see in addition that the estimated coefficient values reported in the first column of these tables are very close to their theoretical values as given in the bias and MSE expansions obtained under Approximation Scheme 1. More precisely, going from top to bottom in column 1 of Table 1 and ignoring the intercept term, we see that the estimated coefficient values are 1, -2.244, 4.561, and -2.320; whereas the corresponding theoretical values, as given in expression (16) are (in the same order) 1, -2, 4, and -2. Similarly, for the MSE, the estimated coefficient values, as reported in column 1 of Table 2, are 0.996, 1.285, -6.370, 11.13, and -5.979 from top to bottom; whereas the corresponding theoretical values as given by expression (18) are 1, 1, -7, 12, and -6. We see this as additional evidence of the appropriateness of the approximations we give in Theorem 4.1. As a further observation on Tables 1 and 2, we note that we have also reported regression results for the cases where only the lead term from Approximation Scheme 1 is included as a regressor (i.e., the case where all regressors involving the factor k_{21}^{-1} are dropped from the regression). Columns 2 and 5 of Table 1 give these results for the bias and the relative bias while the same two columns in Table 2 give them for the MSE and relative MSE. Note that the $\overline{R}^2$ measures recorded for these more parsimonious regressions are in the range 0.9859 to 0.9996. Comparing these numbers with the reported $\overline{R}^2$ for columns 1 and 4 of Tables 1 and 2, we see that the regressions based only on the lead terms fit variations in the bias and MSE functions only slightly worse than regressions which involve second order terms, indicating that in many cases a lead term approximation may be quite

[13] More specifically, we modify regressions (44) and (46) by dividing through by the OLS bias (i.e., $b_{\widehat{\beta}_{OLS}} = \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho = -\beta$ under the canonical model), and we modify regressions (45) and (47) by dividing through by the OLS MSE (i.e., $m_{\widehat{\beta}_{OLS}} = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 = \beta^2$ under the canonical model).

adequate[14].

Finally, in Figures 3 and 4 (in the back of the paper), we have plotted, for a few selected cases, the bias and MSE functions and their approximations under Schemes 1 and 2. More specifically, each diagram in Figure 3 plots the actual and the approximate ABIAS (or AMSE) as a function of k_{21} for s $(= (1+\tau^2)^{-1})$ fixed while the diagrams in Figure 4 plot them as functions of *decreasing* values of s for k_{21} fixed. Note also that, in each diagram, the actual ABIAS (or the actual AMSE) is always represented by the solid line, the approximation under Scheme 1 is given by the shorter dashed line, and the approximation under Scheme 2 is given by the longer dashed line. These diagrams show that whereas the shorter dashed-line graphs based on Approximation Scheme 1 is in most cases virtually indistinguishable from the solid line representing the graphs of the actual ABIAS (or AMSE), the location of the longer dashed lines indicate that the use of Approximation Scheme 2 will lead to substantial error in approximating the ABIAS and the AMSE functions. Moreover, the MSE approximation based on Approximation Scheme 2 is often unreasonable, as its value is negative in many cases. Indeed, Figure 4 suggests that Approximation Scheme 2 only seems to work well when value of s is close to 1 (or when τ^2 is close to 0).

7 Monte Carlo Evidence

In this section, we report the results of a small Monte Carlo study of the sampling behavior of the bias adjusted estimators introduced in Section 5.2. Our experimental setup is based on a special case of the SEM given by equations (1) and (2), with $\gamma = 0$ and $\Phi = 0$. We can write this model in terms of the $t-th$ observation as

$$y_{1t} = y_{2t}\beta + u_t, \tag{48}$$

$$y_{2t} = Z_t'\Pi + v_t, \tag{49}$$

where $t = 1, ..., T$ and where the definitions of y_{1t} (1×1), y_{2t} (1×1), Z_t $(k_2\times 1)$, u_t (1×1), and v_t (1×1) are obvious given the discussion in Section 2. Note that equation (49) is already written in its reduced form since Z_t is presumed to be exogenous, and, as in Remark 3.3(vi), we can also write equation (48) in terms of its reduced form representation as

$$y_{1t} = Z_t'\Gamma_1 + \varepsilon_{1t}, \quad t = 1, ..., T. \tag{50}$$

Data for our Monte Carlo experiments are generated using a canonical version of the model described in expressions (48)-(50) above. In particular, set $\varepsilon_{2t} = v_t$ for all t, and we assume that $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})' \equiv i.i.d.N(0, I_2)$. Moreover, our experimental design is motivated by empirical situations where the number of available instruments is large, so we set $k_2 = 50$ and $T = 500$. We also let $\Pi = (\pi_1,\ \pi_2, ..., \pi_{50})' = (\overline{\pi}, \overline{\pi}, ..., \overline{\pi})'$, so that the degree of relevance of each instrument is assumed to be the same and is given by the magnitude of the scalar parameter $\overline{\pi}$; and, for our experiments, we allow the value of $\overline{\pi}$ to vary over the set $\{0.05, 0.075, 0.10, 0.125, 0.15\}$. The exogenous instruments Z_t are generated by assuming that $Z_t \equiv i.i.d.N(0, I_{k_2})$, and all k_2 instruments are used in the construction of the bias-corrected estimators and of the

[14] Note that in results not reported here, we have found, not surprisingly, that the gain in approximation accuracy from incorporating second order terms tend to be more substantial for smaller k_{21}.

(uncorrected) IV estimator studied in our experiments. Finally, define $w_t = (u_t, v_t)'$ with covariance matrix $E(w_t w_t') = \Sigma = \begin{pmatrix} \sigma_{uu} & \sigma_{vu} \\ \sigma_{vu} & \sigma_{vv} \end{pmatrix}$, and note that our canonical model specification implies that $\sigma_{uu} = 1 + \beta^2$, $\sigma_{uv} = -\beta$, and $\sigma_{vv} = 1$. It follows, of course, that the degree of endogeneity in our data generating processes is determined by the value of the parameter β. Thus, for our experiments, we allow the value of β to vary over the set $\{-0.5, -1.0, -1.5, ..., -5\}$. Note further that, as argued earlier in Section 6, we do not report experimental results for positive values of β because the sign of β only affects the sign of the bias of the IV estimator and has no other material effect on the experimental results.

Table 3 reports the sample bias averaged across 5000 Monte Carlo simulations for the OLS estimator $\widehat{\beta}_{OLS}$ (3rd column), the IV estimator $\widehat{\beta}_{IV}$ (4th column), and our bias adjusted estimators $\widetilde{\beta}_{OLS,i}$, $\widetilde{\beta}_{IV}$, and $\widetilde{\widetilde{\beta}}_{IV,i}$, $i = 1, 2$ (columns 5-9). Looking at Table 3, we note that $\widetilde{\beta}_{IV}$ ranks first as the least biased estimator while $\widetilde{\widetilde{\beta}}_{IV,1}$ places second in each of the 50 experiments based on the different $(\beta, \overline{\pi})$ configurations[15]. At the other end of the spectrum, we see that the two unadjusted estimators, $\widehat{\beta}_{IV}$ and $\widehat{\beta}_{OLS}$, are both amongst the three worst performers in every experimental setting. In particular, the unadjusted OLS estimator $\widehat{\beta}_{OLS}$ is the worst performer in 40 of the 50 experiments conducted, while in each of the remaining 10 experiments, it ranks sixth amongst the group of seven estimators. The unadjusted IV estimator $\widehat{\beta}_{IV}$ fares only slightly better, placing sixth in 20 of the 50 experiments and fifth in the remaining experiments. Overall, the results in Table 3 seem quite encouraging for our bias-adjusted estimators as a group, and particularly for $\widetilde{\beta}_{IV}$ and $\widetilde{\widetilde{\beta}}_{IV,1}$. Indeed, amongst all the bias-adjusted estimators, only $\widetilde{\beta}_{OLS,1}$ does not consistently outperform both $\widehat{\beta}_{IV}$ and $\widehat{\beta}_{OLS}$. Moreover, in many cases, very substantial reduction in bias is achieved when $\widetilde{\beta}_{IV}$ and $\widetilde{\widetilde{\beta}}_{IV,1}$ are used in lieu of $\widehat{\beta}_{IV}$ and $\widehat{\beta}_{OLS}$. For example, in the case where $\beta = -0.5$ and $\overline{\pi} = 0.05$, the reported sample bias of $\widehat{\beta}_{OLS}$ is 0.4441; the reported sample bias of $\widehat{\beta}_{IV}$, on the other hand, is 50% of this OLS bias while the reported sample bias of $\widetilde{\beta}_{IV}$ and $\widetilde{\widetilde{\beta}}_{IV,1}$ are, respectively, only 4.1% and 4.7% of the OLS bias.

The discussion, thus far, has focused on the bias. However, in constructing bias-adjusted estimators, there is always the concern that bias correction may come at the cost of a subtantial increase in the variance of the estimator so that, in the end, the MSE of the bias-corrected estimator is actually higher than that of the unadjusted estimator. To address this concern, we also report in Table 4 the sample MSE of our estimators, averaged across the same 5000 Monte Carlo trials used to compute the corresponding sample bias in Table 3. Table 4, which is organized in the same manner as Table 3, shows clearly that the reduction

[15] Interestingly, $\widetilde{\beta}_{IV}$ outperforms $\widetilde{\widetilde{\beta}}_{IV,1}$ (in the sense of having a slightly lower bias in each experiment), even though bias correction for $\widetilde{\beta}_{IV}$ is based only on the lead term of the bias expansion (16) while bias-correction for $\widetilde{\widetilde{\beta}}_{IV,1}$ is based on a second order term of the expansion in addition to the lead term. An important reason for this, we believe, has to do with the fact that the terms of the bias expansion involve unknown parameters which must be estimated. In particular, we note that, in estimating the bias terms, we have estimated the parameter τ^2 using the Wald statistic $W_{k_{21},T}$. It is easy to see that $W_{k_{21},T} = \tau^2 + Op\left(k_{21}^{-\frac{1}{2}}\right)$, so that the error in estimating τ^2 using $W_{k_{21},T}$ is of order $Op\left(k_{21}^{-\frac{1}{2}}\right)$. It follows that the error of our bias-correction procedure will be of order $Op\left(k_{21}^{-\frac{1}{2}}\right)$ regardless of whether we use one or two terms in the expansion, as it is dominated by the estimation error. As a consequence, we should not be surprised to see both $\widetilde{\beta}_{IV}$ and $\widetilde{\widetilde{\beta}}_{IV,1}$ exhibiting bias of roughly the same order of magnitude, with one performing perhaps slightly better than the other in practice; as it seems to be the case in the results reported in Table 3. On the other hand, in order to be able to take full advantage of the second order term of the bias expansion, we must at least be able to estimate τ^2 such that the error of estimation is of order $O\left(k_{21}^{-(1+\delta)}\right)$ for some $\delta > 0$. We intend to explore this possibility in future research.

in bias achieved by the bias-adjusted IV estimators, $\widetilde{\beta}_{IV}$ and $\widetilde{\widetilde{\beta}}_{IV,1}$, does not come at the expense of higher MSE. Indeed, $\widetilde{\beta}_{IV}$ and $\widetilde{\widetilde{\beta}}_{IV,1}$ are again the top two performers, having the two lowest MSE's in 48 of the 50 experiments reported in Table 4. In fact, only in the case where $\beta = -0.5$ and $\overline{\pi} = 0.05$ and in the case where $\beta = -1.0$ and $\overline{\pi} = 0.05$ did $\widetilde{\beta}_{IV}$ and $\widetilde{\widetilde{\beta}}_{IV,1}$ not place in the top two in Table 4. For these two parameter settings, $\widetilde{\beta}_{OLS,2}$ and $\widetilde{\widetilde{\beta}}_{IV,2}$ were the top two estimators in terms of MSE. This can be explained by the fact that $\widetilde{\beta}_{OLS,2}$ and $\widetilde{\widetilde{\beta}}_{IV,2}$ have been constructed using estimators which restricts Π to be a zero vector. Hence, these estimators tend to perform well in situations where the components of Π (i.e., $\overline{\pi}$) are small and where the level of endogeneity, as measured by the absolute magnitude of β, is small; for, in these cases, the efficiency gained by imposing the zero restriction on Π greatly exceeds the bias induced from misspecifying the model (i.e., the bias which results from shrinking the value of Π to zero when it is in fact only close to zero). Comparing the bias-adjusted estimators with the unadjusted estimators, we see that with the exception of $\widetilde{\beta}_{OLS,1}$, every one of the bias-adjusted estimators outperforms the two unadjusted estimators, $\widehat{\beta}_{IV}$ and $\widehat{\beta}_{OLS}$, in every experiment reported in Table 4. Indeed, across the different experiments, the estimators' ranking based on MSE is very similar to their ranking based on bias; with $\widetilde{\beta}_{OLS,1}$, $\widehat{\beta}_{IV}$, and $\widehat{\beta}_{OLS}$ being the three worst estimators under both criteria.

In sum, the simulation results presented in Tables 3 and 4 show that the bias-adjusted IV estimators, $\widetilde{\beta}_{IV}$ and $\widetilde{\widetilde{\beta}}_{IV,1}$, often offer improvements over the unadjusted estimators, $\widehat{\beta}_{IV}$ and $\widehat{\beta}_{OLS}$, in terms of bias and MSE. In regard to the practical implications of our results, we note that, of course, one would like to use good instruments whenever possible. However, if high quality instruments are not available and one is faced with a situation where one only has available a large number of instruments whose correlation with the endogenous explanatory variable is weak; then, consistent with our theoretical results as reported in Section 5, our Monte Carlo results also suggest that it may be worthwhile, from a point estimation perspective, to use a large number of these weak instruments in constructing bias-corrected estimators.

8 Concluding Remarks

In this paper, we have derived explicit formulae for the asymptotic bias and mean square error of the IV estimator under weak instruments. In addition, we derive approximations for these formulae based on an asymptotic scheme, whereby the value of the concentration parameter and the number of instruments are both taken to infinity while the ratio of the two is assumed to approach a finite limit. These approximations are shown via a series of numerical computations to be quite accurate. Our results allow us to characterize the properties of the (asymptotic) bias and MSE functions of the IV estimator with regard to how they vary with respect to k_{21} and τ^2. 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 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. This result suggests that our bias-corrected estimators may also be useful in standard contexts in which instruments are not weak, although this conjecture is not explored. A series of Monte Carlo experiments documents gains when our bias adjusted estimators are used instead of standard IV and OLS estimators.

This paper is meant as a starting point. We believe that the large k_{21}, local-to-zero framework is potentially useful for addressing a variety of interesting questions related to the weak instrument literature. Future directions for research include the development of methods for the assessment of instrument relevance and for instrument selection based on estimators of the bias and MSE presented in this paper.

Appendix A

This appendix collects a number of lemmas which we will use to establish the main results of our paper. Before presenting the lemmas, however, we first introduce some notations which will appear in the statement and proof of some of our lemmas and theorems. To begin, 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 $Z_{v,1} = \Omega_{11}^{-\frac{1}{2}\prime}(\psi_{Z_1 v} - Q_{Z_1 X}Q_{XX}^{-1}\psi_{Xv})\sigma_{vv}^{-\frac{1}{2}}$, and note that

$$\begin{pmatrix} Z_{u,1} \\ Z_{v,1} \end{pmatrix} \sim N\left(0, \left(\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \otimes I_{k_{21}}\right)\right). \tag{51}$$

In addition, define

$$\begin{aligned} v_1(\mu'\mu, k_{21}) &= (\mu + Z_{v,1})'(\mu + Z_{v,1}) \\ &= \sum\nolimits_{i=1}^{k_{21}} \left(\mu_i + Z_{v,1}^i\right)^2, \qquad (52) \\ v_2(\mu'\mu, k_{21}) &= (\mu + Z_{v,1})' Z_{u,1} \\ &= \sum\nolimits_{i=1}^{k_{21}} \left(\mu_i + Z_{v,1}^i\right) Z_{u,1}^i, \qquad (53) \end{aligned}$$

where μ_i, $Z_{u,1}^i$, and $Z_{v,1}^i$ are the $i-th$ component of μ, $Z_{u,1}$, and $Z_{v,1}$, respectively. Note that we have written $v_1(.,.)$ as a function of $\mu'\mu$ and not μ because v_1 is a noncentral 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 notations, 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}.

The following lemmas will be used in Appendix B to establish the main results of our paper:

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 $T \to \infty$

$$\widehat{\beta}_{IV,T} - \beta_0 \Longrightarrow \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} v_1^{-1} v_2. \tag{54}$$

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: If $x > 0$ and $a, c > 0$, then as $x \to \infty$,

$${}_1F_1(a;c;x) = \frac{\Gamma(c)}{\Gamma(a)} e^x x^{-(c-a)} \left[\sum_{j=0}^{p-1} \frac{(c-a)_j (1-a)_j}{j!} x^{-j} + O(|x|^{-p})\right]. \tag{55}$$

Proof: See Lebedev (1972), pp. 268-271.

Lemma A3: Suppose x is bounded and suppose $a, c \to \infty$ such that $\lim_{a,c\to\infty} \frac{(c-a)x}{c} = 0$. Then,

$${}_1F_1(a;c;x) = e^x \left[\sum_{j=0}^{p-1} \frac{(c-a)_j (-x)^j}{(c)_j \; j!} + O(|c|^{-p})\right]. \tag{56}$$

Proof: The proof follows from Kummer's transform. See, for example, Slater (1960), pp.12, 65-66.

Lemma A4: Let $\chi^2_q(\mu'\mu)$ denote a non-central chi-square random variable with noncentrality parameter $\mu'\mu$ and q degrees of freedom. Also let r denote a positive integer such that $q > 2p$. Then,

$$\begin{aligned} E\left[\left(\chi^2_q(\mu'\mu)\right)^{-p}\right] &= 2^{-p}e^{-\frac{\mu'\mu}{2}}\sum_{j=0}^{\infty}\frac{(\frac{\mu'\mu}{2})^j}{j!}\frac{\Gamma\left(\frac{1}{2}(q+2j)-p\right)}{\Gamma\left(\frac{1}{2}(q+2j)\right)} \\ &= 2^{-p}e^{-\frac{\mu'\mu}{2}}\frac{\Gamma\left(\frac{q}{2}-p\right)}{\Gamma(\frac{q}{2})}\,{}_1F_1\left(\frac{1}{2}q-p;\frac{1}{2}q;\frac{\mu'\mu}{2}\right). \end{aligned} \tag{57}$$

Proof: See Ullah (1974), pp. 145-148.

Lemma A5: If the $(J \times 1)$ vector w is distributed normally with mean vector θ and covariance matrix I_J and suppose $\phi(\cdot)$ is a Borel measurable function. Then,

$$E\left[\phi(w'w)w\right] = \theta E\left[\phi\left(\chi^2_{J+2}(\theta'\theta)\right)\right]. \tag{58}$$

Proof: See Judge and Bock (1978), Theorem 1 of Appendix B.2, pp.321-322.

Lemma A6: If the $(J \times 1)$ random vector w is distributed normally with mean vector θ and covariance matrix I_J and suppose $\phi(\cdot)$ is a Borel measurable function. Then,

$$E\left[\phi(w'w)ww'\right] = E\left[\phi\left(\chi^2_{J+2}(\theta'\theta)\right)\right]I_J + E\left[\phi\left(\chi^2_{J+4}(\theta'\theta)\right)\right]\theta\theta'. \tag{59}$$

Proof: See Judge and Bock (1978), Theorem 3 of Appendix B.2, pp. 323.

Lemma A7: 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)

$$\begin{aligned} &{}_1F_1\left(k_{21}/2-1;\ k_{21}/2;\ \mu'\mu/2\right)\exp\left\{-\left(\mu'\mu/2\right)\right\} \\ =\ & {}_1F_1\left(k_{21}/2-1;\ k_{21}/2;\ \mu'\mu(\tau^2,k_{21})/2\right)\exp\left\{-\left(\mu'\mu(\tau^2,k_{21})/2\right)\right\} \\ =\ & \left(1+\tau^2\right)^{-1} - k_{21}^{-1}\left(1+\tau^2\right)^{-1}\left[2-4\left(1+\tau^2\right)^{-1}+2\left(1+\tau^2\right)^{-2}\right] \\ & -k_{21}^{-2}\left(1+\tau^2\right)^{-2}\left[8-28\left(1+\tau^2\right)^{-1}+32\left(1+\tau^2\right)^{-2}-12\left(1+\tau^2\right)^{-3}\right] \\ & -R^*(k_{21})k_{21}^{-1}\left(1+\tau^2\right)^{-2}+O\left(k_{21}^{-3}\right), \end{aligned} \tag{60}$$

(b)

$$\begin{aligned} &{}_1F_1\left(k_{21}/2-2;\ k_{21}/2-1;\ \mu'\mu/2\right)\exp\left\{-\left(\mu'\mu/2\right)\right\} \\ =\ & {}_1F_1\left(k_{21}/2-2;\ k_{21}/2-1;\ \mu'\mu(\tau^2,k_{21})/2\right)\exp\left\{-\left(\mu'\mu(\tau^2,k_{21})/2\right)\right\} \\ =\ & \left(1+\tau^2\right)^{-1} - k_{21}^{-1}\left(1+\tau^2\right)^{-1}\left[4-6\left(1+\tau^2\right)^{-1}+2\left(1+\tau^2\right)^{-2}\right] \\ & -k_{21}^{-2}\left(1+\tau^2\right)^{-2}\left[24-56\left(1+\tau^2\right)^{-1}+44\left(1+\tau^2\right)^{-2}-12\left(1+\tau^2\right)^{-3}\right] \\ & -R^*(k_{21})k_{21}^{-1}\left(1+\tau^2\right)^{-2}+O\left(k_{21}^{-3}\right). \end{aligned} \tag{61}$$

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

$$
\begin{aligned}
& {}_1F_1\left(k_{21}/2-1;\ k_{21}/2;\ \mu'\mu(\tau^2,k_{21})/2\right)\exp\left\{-\left(\mu'\mu(\tau^2,k_{21})/2\right)\right\} \\
= \ & \frac{\Gamma\left(k_{21}/2\right)}{\Gamma\left(k_{21}/2-1\right)}\int_0^1 \exp\left\{0.5\tau^2 k_{21}(t-1)\right\}\exp\left\{0.5R^*(k_{21})(t-1)\right\} t^{(k_{21}-4)/2}dt \\
= \ & \left[0.5\left(k_{21}-2\right)\right]\int_0^1 \exp\left\{k_{21}h_1(t)\right\}\exp\left\{0.5R^*(k_{21})(t-1)\right\}dt, \qquad (62)
\end{aligned}
$$

where $h_1(t) = 0.5\left[\tau^2(t-1)+\log t\right] - (2/k_{21})\log t$. Given the integral representation (62), we can obtain the expansion given by the right-hand side of expression (60) by applying a Laplace approximation to this integral representation. We note that the maximum of the integrand of (62) 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, we first split up this integral as follows:

$$
\begin{aligned}
& \left[0.5\left(k_{21}-2\right)\right]\int_0^1 \exp\left\{k_{21}h_1(t)\right\}\exp\left\{0.5R^*(k_{21})(t-1)\right\}dt \\
= \ & \left[0.5\left(k_{21}-2\right)\right]\int_{1-1/\sqrt{k_{21}}}^1 \exp\left\{k_{21}h_1(t)\right\}\exp\left\{0.5R^*(k_{21})(t-1)\right\}dt \\
& +\left[0.5\left(k_{21}-2\right)\right]\int_0^{1-1/\sqrt{k_{21}}} \exp\left\{k_{21}h_1(t)\right\}\exp\left\{0.5R^*(k_{21})(t-1)\right\}dt \\
= \ & I_1 + I_2 \quad (say), \qquad (63)
\end{aligned}
$$

We shall handle I_2 first. Note that

$$
\begin{aligned}
& \left[0.5\left(k_{21}-2\right)\right]\int_0^{1-1/\sqrt{k_{21}}} \exp\left\{k_{21}h_1(t)\right\}\exp\left\{0.5R^*(k_{21})(t-1)\right\}dt \\
\leq \ & \left[0.5\left(k_{21}-2\right)\right]\exp\left\{-\left(0.5\tau^2\sqrt{k_{21}}\right)\right\}\left(1-k_{21}^{-\frac{1}{2}}\right)^{(k_{21}-2)/2}\exp\left\{-0.5k_{21}^{-\frac{1}{2}}R^*(k_{21})\right\} \\
= \ & O\left(k_{21}\exp\left\{-\left(0.5\tau^2\sqrt{k_{21}}\right)\right\}\left(1-k_{21}^{-\frac{1}{2}}\right)^{(k_{21}-2)/2}\right), \qquad (64)
\end{aligned}
$$

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 $I_1 = \left[0.5\left(k_{21}-2\right)\right]\int_{-1/\sqrt{k_{21}}}^0 \exp\left\{k_{21}h_2(r)\right\}\exp\left\{0.5R^*(k_{21})r\right\}dr$ 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_1in 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

$$
h_2(r) = h_2(0)+h_2'(0)r+\left(h_2^{(2)}(0)/2!\right)r^2+\left(h_2^{(3)}(0)/3!\right)r^3+\left(h_2^{(4)}(r^*)/4!\right)r^4, \qquad (65)
$$

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 $-1/\sqrt{k_{21}} \le r \le 0$,

$$\begin{aligned} \left| h_2(r) - \sum\nolimits_{i=1}^{3} \frac{h_2^{(i)}(0)}{i!} r^i \right| &= \left| \frac{h_2^{(4)}(r^*)}{4!} r^4 \right| \\ &\le \left[M(k_{21}) r^4 \right] / 4!. \end{aligned} \tag{66}$$

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

$$\begin{aligned} &\left(\frac{k_{21}-2}{2}\right) \int_{-\frac{1}{\sqrt{k_{21}}}}^{0} \exp\left\{ k_{21} \left(\sum_{i=1}^{3} \frac{h_2^{(i)}(0)}{i!} r^i - \frac{M(k_{21})r^4}{4!} \right) \right\} \exp\left\{ \frac{r}{2} R^*(k_{21}) \right\} dr \\ \le &\left(\frac{k_{21}-2}{2}\right) \int_{-1/\sqrt{k_{21}}}^{0} \exp\left\{ k_{21} h_2(r) \right\} \exp\left\{ 0.5 r R^*(k_{21}) \right\} dr \\ \le &\left(\frac{k_{21}-2}{2}\right) \int_{-\frac{1}{\sqrt{k_{21}}}}^{0} \exp\left\{ k_{21} \left(\sum_{i=1}^{3} \frac{h_2^{(i)}(0)}{i!} r^i + \frac{M(k_{21})r^4}{4!} \right) \right\} \exp\left\{ \frac{r}{2} R^*(k_{21}) \right\} dr. \end{aligned} \tag{67}$$

Let I_3 denote the upper bound integral in expression (67). To evaluate I_3, we rewrite it as

$$\begin{aligned} I_3 = &\left(\frac{k_{21}-2}{2}\right) \int_{-\frac{1}{\sqrt{k_{21}}}}^{0} \exp\left\{ k_{21} h_2'(0) r \right\} \exp\left\{ k_{21} \left(\sum_{i=2}^{3} \frac{h_2^{(i)}(0)}{i!} r^i + \frac{M(k_{21})r^4}{4!} \right) \right\} \\ &\exp\left\{ \frac{r}{2} R^*(k_{21}) \right\} dr. \end{aligned} \tag{68}$$

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,

$$\begin{aligned} I_3 = &\left(0.5\left(k_{21}-2\right)\right) \left[\int_{-1/\sqrt{k_{21}}}^{0} \exp\left\{ k_{21} h_2'(0) r \right\} \left(1 + \left[k_{21} h_2^{(2)}(0)/2! \right] r^2 \right. \right. \\ &\left. \left. + \left[k_{21} h_2^{(3)}(0)/3! \right] r^3 + \left[k_{21} \left(h_2^{(2)}(0) \right)^2 (2!)^{-3} \right] r^4 + 0.5 R^*(k_{21}) r \right) dr + O\left(k_{21}^{-4}\right) \right] \end{aligned} \tag{69}$$

$$\begin{aligned} = &\left(0.5\left(k_{21}-2\right)\right) \left[2k_{21}^{-1}\left(1+\tau^2\right)^{-1} + 8k_{21}^{-2}\left(1+\tau^2\right)^{-2} + 32k_{21}^{-3}\left(1+\tau^2\right)^{-3} \right. \\ &-4k_{21}^{-2}\left(1+\tau^2\right)^{-3} + 16k_{21}^{-3}\left(1+\tau^2\right)^{-3} - 48k_{21}^{-3}\left(1+\tau^2\right)^{-4} \\ &\left. -16k_{21}^{-3}\left(1+\tau^2\right)^{-4} + 24k_{21}^{-3}\left(1+\tau^2\right)^{-5} - 2R^*(k_{21})k_{21}^{-2}\left(1+\tau^2\right)^{-2} + O\left(k_{21}^{-4}\right) \right] \end{aligned} \tag{70}$$

$$\begin{aligned} = &\left(1+\tau^2\right)^{-1} - k_{21}^{-1}\left(1+\tau^2\right)^{-1} \left[2 - 4\left(1+\tau^2\right)^{-1} + 2\left(1+\tau^2\right)^{-2} \right] - k_{21}^{-2}\left(1+\tau^2\right)^{-2} [8- \\ &28\left(1+\tau^2\right)^{-1} + 32\left(1+\tau^2\right)^{-2} - 12\left(1+\tau^2\right)^{-3}] - R^*(k_{21})k_{21}^{-1}\left(1+\tau^2\right)^{-3} + O\left(k_{21}^{-3}\right). \end{aligned} \tag{71}$$

By a similar argument, it can be shown that the lower bound integral in expression (67) can also be approximated by the right-hand side of expression (71). It, thus, follows that

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

$$= \left(1+\tau^2\right)^{-1}-k_{21}^{-1}\left(1+\tau^2\right)^{-1}\left[2-4\left(1+\tau^2\right)^{-1}+2\left(1+\tau^2\right)^{-2}\right]-k_{21}^{-2}\left(1+\tau^2\right)^{-2}[8-$$
$$28\left(1+\tau^2\right)^{-1}+32\left(1+\tau^2\right)^{-2}-12\left(1+\tau^2\right)^{-3}\Big]-R^*(k_{21})k_{21}^{-1}\left(1+\tau^2\right)^{-3}+O\left(k_{21}^{-3}\right) \quad (72)$$

Finally, the result given in part (a) follows immediately from expressions (64) and (72).

Lemma A8: Suppose that (1), (2) and Assumptions 1 and 2 hold. Then, the following convergence results hold jointly as $T\to\infty$:

(a) $(u'M_Xu/T,\, y_2'M_Xu/T,\, y_2'M_Xy_2/T)\xrightarrow{p}(\sigma_{uu},\,\sigma_{uv},\,\sigma_{vv})$.

(b) $Z_1'M_XZ_1/T\xrightarrow{p}\Omega_{11}$, where $\Omega_{11}=Q_{Z_1Z_1}-Q_{Z_1X}Q_{XX}^{-1}Q_{XZ_1}$.

(c) $\left\{(Z_1'M_XZ_1)^{-\frac{1}{2}}Z_1'M_Xu,(Z_1'M_XZ_1)^{-\frac{1}{2}}Z_1'M_Xv\right\}\Longrightarrow\left\{Z_{u,1}\sigma_{uu}^{\frac{1}{2}},Z_{v,1}\sigma_{vv}^{\frac{1}{2}}\right\}$,
where $(Z_{u,1}',\,Z_{v,1}')'$ has joint normal distribution given by (51).

(d) $(Z_1'M_XZ_1/T)^{-\frac{1}{2}}\left(Z_1'M_Xy_2/\sqrt{T}\right)\Longrightarrow(\mu+Z_{v,1})\,\sigma_{vv}^{\frac{1}{2}}$.

(e) $\big(y_2'M_XZ_1(Z_1'M_XZ_1)^{-1}Z_1'M_Xu,\; y_2'M_XZ_1(Z_1'M_XZ_1)^{-1}Z_1'M_Xy_2,$
$u'M_XZ_1(Z_1'M_XZ_1)^{-1}Z_1'M_Xu\big)\Longrightarrow\left(\sigma_{vv}^{\frac{1}{2}}v_2\sigma_{uu}^{\frac{1}{2}},\;\sigma_{vv}v_1,\;\sigma_{uu}Z_{u,1}'Z_{u,1}\right)$.

(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: Part (a) is identical to part (a) of Lemma A1 of Staiger and Stock (1997) and is proved there. Parts (b)-(e) are similar to parts (b)-(e) of Lemma A1 of Staiger and Stock (1997), the only difference being that Lemma A1 of Staiger and Stock (1997) gives convergence results for sample moments involving the entire instrument matrix Z whereas our lemma here involves Z_1, the submatrix of Z obtained via column selection. Hence, parts (b)-(e) can be proved by minor modifications of the proof of parts (b)-(e) of Lemma A1 of Staiger and Stock (1997).

To show part (f), write
$\frac{u'M_{(Z,\,X)}u}{T}=\frac{u'M_Xu}{T}-\frac{u'M_XZ}{T}\left(\frac{Z'M_XZ}{T}\right)^{-1}\frac{Z'M_Xu}{T}$, $\frac{y_2'M_{(Z,\,X)}u}{T}=\frac{y_2'M_Xu}{T}-\frac{y_2'M_XZ}{T}\left(\frac{Z'M_XZ}{T}\right)^{-1}\frac{Z'M_Xu}{T}$, and
$\frac{y_2'M_{(Z,\,X)}y_2}{T}=\frac{y_2'M_Xy_2}{T}-\frac{y_2'M_XZ}{T}\left(\frac{Z'M_XZ}{T}\right)^{-1}\frac{Z'M_Xy_2}{T}$. Now, part (e) of Lemma A9 of Staiger and Stock (1997) implies that
$\frac{u'M_XZ}{T}\left(\frac{Z'M_XZ}{T}\right)^{-1}\frac{Z'M_Xu}{T}=Op\left(\frac{1}{T}\right)$, $\frac{y_2'M_XZ}{T}\left(\frac{Z'M_XZ}{T}\right)^{-1}\frac{Z'M_Xu}{T}=Op\left(\frac{1}{T}\right)$, and $\frac{y_2'M_XZ}{T}\left(\frac{Z'M_XZ}{T}\right)^{-1}\frac{Z'M_Xy_2}{T}=Op\left(\frac{1}{T}\right)$. The results of part (f) then follow immediately from part (a) of this Lemma and the Slutsky's Theorem.

To show part (g), first note that g_{11} and g_{12} are related to elements of the structural error covariance matrix Σ by the relations: $g_{11}=\sigma_{uu}+2\sigma_{uv}\beta+\sigma_{vv}\beta^2$ and $g_{11}=\sigma_{uu}+\sigma_{vv}\beta$, where $\sigma_{vv}=g_{22}$. Next, observe that $\frac{y_1'M_{(Z,\,X)}y_1}{T}=\frac{u'M_{(Z,\,X)}u}{T}+2\frac{y_2'M_{(Z,\,X)}u}{T}\beta+\frac{y_2'M_{(Z,\,X)}y_2}{T}\beta^2$ and $\frac{y_1'M_{(Z,\,X)}y_2}{T}=\frac{y_2'M_{(Z,\,X)}u}{T}+\frac{y_2'M_{(Z,\,X)}y_2}{T}\beta$. Thus, it follows immediately from part (f) of this Lemma and the Slutsky's Theorem that $\frac{y_1'M_{(Z,\,X)}y_1}{T}\xrightarrow{p}\sigma_{uu}+2\sigma_{uv}\beta+\sigma_{vv}\beta^2=g_{11}$ and $\frac{y_1'M_{(Z,\,X)}y_2}{T}\xrightarrow{p}\sigma_{uu}+\sigma_{vv}\beta=g_{12}$.

Lemma A9: 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 $\mu'\mu=\tau^2k_{21}+O(k_{21}^{-1})=\mu'\mu(\tau^2,k_{21})$. Then, as $k_{21}\to\infty$, the following results hold:

(a) $\frac{v_1(\mu'\mu(\tau^2,\, k_{21}),k_{21})}{k_{21}} \xrightarrow{p} (1+\tau^2)$.

(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 (51), 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 A10: 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 3.1: To show part (a), we note that by Lemma A1, $U_T = \widehat{\beta}_{IV,T} - \beta_0 \Rightarrow \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} v_1^{-1}v_2 \equiv U$ (say). Moreover, given Assumption 3, we have by Theorem 5.4 of Billingsley (1968) that $\lim_{T\to\infty} E(U_T) = \lim_{T\to\infty} E\left[\widehat{\beta}_{IV,T} - \beta_0\right] = E\left[\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} v_1^{-1}v_2\right] = E(U)$. It follows that to derive the asymptotic bias of $\widehat{\beta}_{IV}$, we need merely to give an explicit form for $E\left[\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} v_1^{-1}v_2\right]$. To proceed, note that, given (51), we can write $Z_{u,1} = Z_{v,1}\rho + Z_{u1.v1}$, where $Z_{u1.v1} \sim N(0, (1-\rho^2)I_{k_{21}})$ represents the projection error and is, thus, independent of $Z_{v,1}$. Next, we rewrite the limiting random variable U as

$$\begin{aligned} U &= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} v_1^{-1}v_2 = \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} \left[(\mu + Z_{v,1})'(\mu + Z_{v,1})\right]^{-1} (\mu + Z_{v,1})'Z_{u,1} \\ &= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} \left[(\mu + Z_{v,1})'(\mu + Z_{v,1})\right]^{-1} (\mu + Z_{v,1})'(Z_{v,1}\rho + Z_{u1.v1}), \qquad (73) \end{aligned}$$

so that making use of the law of iterated expectations, we have

$$\begin{aligned} E(U) &= E_{Z_{v,1}}\left[E_{Z_{u,1}|Z_{v,1}}\left(\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} \left[(\mu + Z_{v,1})'(\mu + Z_{v,1})\right]^{-1} (\mu + Z_{v,1})'(Z_{v,1}\rho + Z_{u1.v1})\right)\right] \\ &= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} E_{Z_{v,1}}\left[\left[(\mu + Z_{v,1})'(\mu + Z_{v,1})\right]^{-1} (\mu + Z_{v,1})'Z_{v,1}\rho\right], \qquad (74) \end{aligned}$$

where $E_{Z_{v,1}}(\cdot)$ and $E_{Z_{u,1}|Z_{v,1}}(\cdot)$ denote, respectively, the expectation taken with respect to the marginal density of $Z_{v,1}$ and the expectation taken with respect to the conditional density of $Z_{u,1}$ given $Z_{v,1}$. Now, to evaluate the right-hand side of (74), we note that

$$
\begin{aligned}
& \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} E_{Z_{v,1}}\left[\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'Z_{v,1}\rho\right] \\
= \; & \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} E_{Z_{v,1}}\left[\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'(\mu+Z_{v,1}-\mu)\rho\right] \\
= \; & \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho-\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} E_{Z_{v,1}}\left[\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'\mu\rho\right] \\
= \; & \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho-\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\mu'\mu E\left(\left[\chi^2_{k_{21}+2}(\mu'\mu)\right]^{-1}\right). \qquad (75)
\end{aligned}
$$

where the last line of expression (75) follows from Lemma A5 by noting that $(\mu+Z_{v,1})\sim N(\mu,I_{k_{21}})$ and so $(\mu+Z_{v,1})'(\mu+Z_{v,1})\sim\chi^2_{k_{21}}(\mu'\mu)$. Finally, applying Lemma A4 to (75), we obtain

$$
\begin{aligned}
\lim_{T\to\infty} E\left[\widehat{\beta}_{IV,T}-\beta_0\right] &= E(U) \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1-\left(\frac{\mu'\mu}{2}\right)e^{-\frac{\mu'\mu}{2}}\frac{\Gamma(k_{21}/2)}{\Gamma(k_{21}/2+1)}\,{}_1F_1\left(\frac{k_{21}}{2};\frac{k_{21}}{2}+1;\frac{\mu'\mu}{2}\right)\right] \qquad (76) \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\Big[1-e^{-\frac{\mu'\mu}{2}}\left\{\Gamma(k_{21}/2)/\Gamma(k_{21}/2+1)\right\}(k_{21}/2) \\
&\quad\times\left\{{}_1F_1\left(k_{21}/2;\;k_{21}/2;\;\mu'\mu/2\right)-{}_1F_1\left(k_{21}/2-1;\;k_{21}/2;\;\mu'\mu/2\right)\right\}\Big] \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\Big[1-e^{-\frac{\mu'\mu}{2}}\,{}_1F_1\left(k_{21}/2;\;k_{21}/2;\;\mu'\mu/2\right) \\
&\quad+e^{-\frac{\mu'\mu}{2}}\,{}_1F_1\left(k_{21}/2-1;\;k_{21}/2;\;\mu'\mu/2\right)\Big] \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho e^{-\frac{\mu'\mu}{2}}\,{}_1F_1\left(k_{21}/2-1;\;k_{21}/2;\;\mu'\mu/2\right), \qquad (77)
\end{aligned}
$$

where the third equality above follows from the recurrence relation
$z\;{}_1F_1(\alpha+1;\gamma+1;z)=\gamma\times\left[{}_1F_1(\alpha+1;\gamma;z)-{}_1F_1(\alpha;\gamma;z)\right]$
and where the sixth equality above follows from the fact that ${}_1F_1(\alpha;\alpha;z)=e^z$.

To show part (b), note that $\frac{\mu'\mu}{2}>0$, $\frac{k_{21}}{2}>0$, and $\frac{k_{21}}{2}+1>0$. Hence, direct application of Lemma A2 yields $\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho e^{-\frac{\mu'\mu}{2}}\,{}_1F_1\left(k_{21}/2-1;\;k_{21}/2;\mu'\mu/2\right)=$
$=\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho e^{-\frac{\mu'\mu}{2}}\left(\Gamma\left(k_{21}/2\right)/\Gamma(k_{21}/2-1)\right)e^{\frac{\mu'\mu}{2}}\left(\mu'\mu/2\right)^{-1}\left[1+O((\mu'\mu)^{-1})\right]=$
$=\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho(k_{21}-2)\left(\mu'\mu\right)^{-1}\left[1+O((\mu'\mu)^{-1})\right]=O((\mu'\mu)^{-1})$.

To show part (c), note that $\lim_{k_{21}\to\infty}\frac{\left(\frac{1}{2}k_{21}-\frac{1}{2}k_{21}+1\right)\left(\frac{\mu'\mu}{2}\right)}{\left(\frac{1}{2}k_{21}\right)}=\lim_{k_2\to\infty}\left(\frac{\mu'\mu}{k_{21}}\right)=0$. Hence, direct application of Lemma A3 gives $\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho e^{-\frac{\mu'\mu}{2}}\,{}_1F_1\left(k_{21}/2-1;\;k_{21}/2;\;\mu'\mu/2\right)=\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho e^{-\frac{\mu'\mu}{2}}e^{\frac{\mu'\mu}{2}}\left[1+O(k_{21}^{-1})\right]=\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1+O(k_{21}^{-1})\right]\to\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho$ as $k_{21}\to\infty$.

To show (d), note that, based on (7), we can write the bias formula in its infinite series form as follows:

$$
\begin{aligned}
b_{\widehat{\beta}_{IV}}(\mu'\mu,k_{21}) &= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho e^{-\frac{\mu'\mu}{2}}\left[\sum_{j=0}^{\infty}\frac{(k_{21}/2-1)_j}{(k_{21}/2)j}\frac{(\mu'\mu/2)^j}{j!}\right] \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho e^{-\frac{\mu'\mu}{2}}\left[\sum_{j=0}^{\infty}\left(\frac{k_{21}-2}{k_{21}+2j-2}\right)\frac{(\mu'\mu/2)^j}{j!}\right] \qquad (78)
\end{aligned}
$$

.

Let $f(\mu'\mu, k_{21}) = e^{-\frac{\mu'\mu}{2}}\left[\sum_{j=0}^{\infty}\left(\frac{k_{21}-2}{k_{21}+2j-2}\right)\frac{(\mu'\mu/2)^j}{j!}\right]$, and note further that $f(\mu'\mu = 0, k_{21}) = 1$. Also, from the proof of part (b) above, we know that $\lim_{\mu'\mu\to\infty} f(\mu'\mu, k_{21}) = 0$. Moreover, note that, under the assumption that $k_{21} \geq 4$,

$$\begin{aligned}
\frac{\partial f(\mu'\mu, k_{21})}{\partial(\mu'\mu)} &= \left(\frac{1}{2}\right)e^{-\frac{\mu'\mu}{2}}\sum_{j=0}^{\infty}\left(\frac{k_{21}-2}{k_{21}+2j-2}\right)\frac{j\,(\mu'\mu/2)^{j-1}}{j!} - \sum_{j=0}^{\infty}\left(\frac{k_{21}-2}{k_{21}+2j-2}\right)\frac{(\mu'\mu/2)^j}{j!} \\
&= \left(\frac{1}{2}\right)e^{-\frac{\mu'\mu}{2}}\sum_{j=1}^{\infty}\left(\frac{k_{21}-2}{k_{21}+2j-2}\right)\frac{\left(\frac{\mu'\mu}{2}\right)^{j-1}}{(j-1)!} - \sum_{j=0}^{\infty}\left(\frac{k_{21}-2}{k_{21}+2j-2}\right)\frac{(\mu'\mu/2)^j}{j!} \\
&= \left(\frac{1}{2}\right)e^{-\frac{\mu'\mu}{2}}\sum_{j=0}^{\infty}\frac{(\mu'\mu/2)^j}{j!}\frac{(k_{21}-2)\left[(k_{21}+2j-2)-(k_{21}+2j)\right]}{(k_{21}+2j-2)(k_{21}+2j)} \\
&= -e^{-\frac{\mu'\mu}{2}}\left[\sum_{j=0}^{\infty}\left(\frac{k_{21}-2}{(k_{21}+2j-2)\,(k_{21}+2j)}\right)\frac{(\mu'\mu/2)^j}{j!}\right] < 0,
\end{aligned} \tag{79}$$

where term-by-term differentiation is justified by the absolute and uniform convergence of the infinite series representation of $f(\mu'\mu, k_{21})$ and of the infinite series (79). It follows that $0 \leq f(\mu'\mu, k_{21}) \leq 1$, and is a monotonically decreasing function of $(\mu'\mu)$ for $(\mu'\mu) \in [0, \infty)$. Moreover, from expression (78) and the definition of $f(\mu'\mu, k_{21})$, we see that $|b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21})| = |\rho|\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} f(\mu'\mu, k_{21})$, so that $|b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21})|$ depends on $\mu'\mu$ only through the factor $f(\mu'\mu, k_{21})$. Hence, $|b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21})|$ is a monotonically decreasing function of $\mu'\mu$ for $\mu'\mu \in [0, \infty)$ and $\sigma_{uv} \neq 0$.

To show (e), we differentiate the infinite series representation of $f(\mu'\mu, k_{21})$ term-by-term to obtain

$$\frac{\partial f(\mu'\mu, k_{21})}{\partial k_{21}} = e^{-\frac{\mu'\mu}{2}}\sum_{j=1}^{\infty}\left[\frac{2j}{(k_{21}+2j-2)^2}\right]\frac{(\mu'\mu/2)^j}{j!} > 0, \tag{80}$$

noting that interchanging the operations of differentiation and summation is jusified by the absolute and uniform convergence of the infinite series involved for $k_{21} \geq 4$. It follows that $f(\mu'\mu, k_{21})$ and, thus, $|b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21})|$ are monotonically increasing functions of k_{21} for $\mu'\mu$ fixed and $\sigma_{uv} \neq 0$.

Proof of Theorem 3.2: To show part (a), note that by Theorem 5.4 of Billingsley (1968) and Lemma A1 above, we have that $\lim_{T\to\infty} E\left[\widehat{\beta}_{IV,T} - \beta_0\right]^2 = \lim_{T\to\infty} E(U_T^2) = E(U^2) = E[\sigma_{uu}\sigma_{vv}^{-1}v_1^{-1}v_2^2v_1^{-1}]$. Hence, as with the proof of part (a) of Theorem 3.1, the derivation of the AMSE only entails the derivation of an explicit form for $E[\sigma_{uu}\sigma_{vv}^{-1}v_1^{-1}v_2^2v_1^{-1}]$. To proceed, note that, using expression (51) and the decomposition $Z_{u,1} = Z_{v,1}\rho + Z_{u1.v1}$, we can write $U^2 = \sigma_{uu}\sigma_{vv}^{-1}v_1^{-1}v_2^2v_1^{-1} = \sigma_{uu}\sigma_{vv}^{-1}\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'(Z_{v,1}\rho + Z_{u1.v1})(Z_{v,1}\rho + Z_{u1.v1})'(\mu+Z_{v,1})\times$
$\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}$, so that making use of the law of iterated expectations, we have

$$
\begin{aligned}
E(U^2) &= \sigma_{uu}\sigma_{vv}^{-1}E_{Z_{v,1}}\left[E_{Z_{u,1}|Z_{v,1}}\left(\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'(Z_{v,1}\rho+Z_{u1.v1})\right.\right.\\
&\quad \left.\left.(Z_{v,1}\rho+Z_{u1.v1})'(\mu+Z_{v,1})\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\right)\right]\\
&= \sigma_{uu}\sigma_{vv}^{-1}E_{Z_{v,1}}\left(\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'\left(Z_{v,1}Z_{v,1}'\rho^2+(1-\rho^2)I_{k_{21}}\right)\right.\\
&\quad \left.(\mu+Z_{v,1})\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\right)\\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2E_{Z_{v,1}}\left(\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'Z_{v,1}Z_{v,1}'(\mu+Z_{v,1})\right.\\
&\quad \left.\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\right)+\sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)E_{Z_{v,1}}\left(\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\right)\\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2E_{Z_{v,1}}\left(\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'(\mu+Z_{v,1}-\mu)(\mu+Z_{v,1}-\mu)'\right.\\
&\quad \left.(\mu+Z_{v,1})\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\right)\\
&\quad +\sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)E_{Z_{v,1}}\left(\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\right)\\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left\{1-2E_{Z_{v,1}}\left(\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'\mu\right)\right.\\
&\quad +E_{Z_{v,1}}\left(\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\mu'(\mu+Z_{v,1})(\mu+Z_{v,1})'\mu\right.\\
&\quad \left.\left.\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\right)+\rho^{-2}\left(1-\rho^2\right)E_{Z_{v,1}}\left(\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\right)\right\}, \qquad (81)
\end{aligned}
$$

where $E_{Z_{v,1}}(\cdot)$ and $E_{Z_{u,1}|Z_{v,1}}(\cdot)$ are expectation operators as defined in the proof of part (a) of Theorem 3.1. Now, to evaluate the expression to the right of the last equality sign above, we note that since $(\mu+Z_{v,1})\sim N(\mu,I_{k_{21}})$ and $(\mu+Z_{v,1})'(\mu+Z_{v,1})\sim\chi^2_{k_{21}}(\mu'\mu)$, we can apply Lemmas A5 and A6 to (81) above to obtain

$$
\begin{aligned}
E(U^2) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left\{1-2E\left(\left(\chi^2_{k_{21}+2}(\mu'\mu)\right)^{-1}\right)\mu'\mu+\mu'\left[E\left(\left(\chi^2_{k_{21}+2}(\mu'\mu)\right)^{-2}\right)I_{k_{21}}\right.\right.\\
&\quad \left.\left.+\left(E\left(\chi^2_{k_{21}+4}(\mu'\mu)\right)^{-2}\right)\mu\mu'\right]\mu+\rho^{-2}\left(1-\rho^2\right)E\left(\left(\chi^2_{k_{21}}(\mu'\mu)\right)^{-1}\right)\right\}. \qquad (82)
\end{aligned}
$$

Finally, applying Lemma A4 to expression (82) above, we obtain

$$
\begin{aligned}
\lim_{T\to\infty}E\left[\widehat{\beta}_{IV,T}-\beta_0\right]^2 &= E(U^2)\\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\\
&\quad \left\{1-(\mu'\mu)\,e^{-\frac{\mu'\mu}{2}}\left[\Gamma(k_{21}/2)/\Gamma(k_{21}/2+1)\right]\ {}_1F_1\left(k_{21}/2;\ k_{21}/2+1;\ \mu'\mu/2\right)\right.\\
&\quad +(\mu'\mu/4)\,e^{-\frac{\mu'\mu}{2}}\left[\Gamma(k_{21}/2-1)/\Gamma(k_{21}/2+1)\right]\ {}_1F_1\left(k_{21}/2-1;k_{21}/2+1;\mu'\mu/2\right)\\
&\quad +(\mu'\mu/2)^2\,e^{-\frac{\mu'\mu}{2}}\left[\Gamma(k_{21}/2)/\Gamma(k_{21}/2+2)\right]\ {}_1F_1\left(k_{21}/2;\ k_{21}/2+2;\ \mu'\mu/2\right)\\
&\quad +\left(2\rho^2\right)^{-1}\left(1-\rho^2\right)e^{-\frac{\mu'\mu}{2}}\left[\Gamma(k_{21}/2-1)/\Gamma(k_{21}/2)\right]\times\\
&\quad \left.{}_1F_1\left(k_{21}/2-1;\ k_{21}/2;\ \mu'\mu/2\right)\right\}\\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2A\ (say). \qquad (83)
\end{aligned}
$$

Now, rewrite A as

$$
\begin{aligned}
A &= 1-\left(\mu'\mu\right)e^{-\left(\mu'\mu/2\right)}2k_{21}^{-1}\ {}_1F_1\left(k_{21}/2;\ k_{21}/2+1;\ \mu'\mu/2\right)\\
&\quad+\left(\mu'\mu/4\right)e^{-\left(\mu'\mu/2\right)}4\left(k_{21}(k_{21}-2)\right)^{-1}\ {}_1F_1\left(k_{21}/2-1;\ k_{21}/2+1;\ \mu'\mu/2\right)\\
&\quad+\left(\mu'\mu/2\right)^2e^{-\left(\mu'\mu/2\right)}4\left(k_{21}(k_{21}+2)\right)^{-1}\ {}_1F_1\left(k_{21}/2;\ k_{21}/2+2;\ \mu'\mu/2\right)\\
&\quad+\left(2\rho^2\right)^{-1}\left(1-\rho^2\right)e^{-\left(\mu'\mu/2\right)}2\left(k_{21}-2\right)^{-1}\ {}_1F_1\left(k_{21}/2-1;\ k_{21}/2;\ \mu'\mu/2\right).
\end{aligned}
\tag{84}
$$

Next, note that successive application of the recurrence relation $z\ {}_1F_1\left(\alpha+1;\gamma+1;z\right)=\gamma\left[{}_1F_1\left(\alpha+1;\gamma;z\right)-\ {}_1F_1\left(\alpha;\gamma;z\right)\right]$ yields

$$
\begin{aligned}
A &= 1-\left[\left(\mu'\mu\right)e^{-\left(\mu'\mu/2\right)}2k_{21}^{-1}\ {}_1F_1\left(k_{21}/2;k_{21}/2+1;\mu'\mu/2\right)\right]+\left[\left(\mu'\mu/4\right)e^{-\left(\mu'\mu/2\right)}4\times\right.\\
&\quad\left.\left(k_{21}(k_{21}-2)\right)^{-1}\ {}_1F_1\left(k_{21}/2-1;k_{21}/2+1;\mu'\mu/2\right)\right]+\left[\left(\mu'\mu/2\right)e^{-\left(\mu'\mu/2\right)}4\left(k_{21}(k_{21}+2)\right)^{-1}\times\right.\\
&\quad\left(\left(k_{21}+2\right)/2\right)\left\{{}_1F_1\left(k_{21}/2;\ k_{21}/2+1;\ \mu'\mu/2\right)-\ {}_1F_1\left(k_{21}/2-1;\ k_{21}/2+1;\ \mu'\mu/2\right)\right\}]\\
&\quad+\left[\left(2\rho^2\right)^{-1}\left(1-\rho^2\right)e^{-\left(\mu'\mu/2\right)}2\left(k_{21}-2\right)^{-1}\ {}_1F_1\left(k_{21}/2-1;k_{21}/2;\mu'\mu/2\right)\right]\\
&= 1-\left[\left(\mu'\mu/2\right)e^{-\left(\mu'\mu/2\right)}2k_{21}^{-1}\ {}_1F_1\left(k_{21}/2;\ k_{21}/2+1;\ \mu'\mu/2\right)\right]+\left[e^{-\left(\mu'\mu/2\right)}2\left(k_{21}(k_{21}-2)\right)^{-1}\times\right.\\
&\quad\left(k_{21}/2\right)\left\{{}_1F_1\left(k_{21}/2-1;\ k_{21}/2;\ \mu'\mu/2\right)-\ {}_1F_1\left(k_{21}/2-2;\ k_{21}/2;\ \mu'\mu/2\right)\right\}]-[\left(\mu'\mu/2\right)\times\\
&\quad\left.e^{-\left(\mu'\mu/2\right)}\left(2/k_{21}\right)\ {}_1F_1\left(k_{21}/2-1;k_{21}/2+1;\mu'\mu/2\right)\right]+\left[\rho^{-2}\left(1-\rho^2\right)e^{-\left(\mu'\mu/2\right)}\left(k_{21}-2\right)^{-1}\times\right.\\
&\quad{}_1F_1\left(k_{21}/2-1;k_{21}/2;\mu'\mu/2\right)]\\
&= 1-\left[e^{-\left(\mu'\mu/2\right)}\left\{{}_1F_1\left(k_{21}/2;k_{21}/2;\mu'\mu/2\right)-\ {}_1F_1\left(k_{21}/2-1;k_{21}/2;\mu'\mu/2\right)\right\}\right]-\left[e^{-\left(\mu'\mu/2\right)}\times\right.\\
&\quad\left.\left(k_{21}-2\right)^{-1}\ {}_1F_1\left(k_{21}/2-2;k_{21}/2;\mu'\mu/2\right)\right]-\left[e^{-\left(\mu'\mu/2\right)}\ \left\{{}_1F_1\left(k_{21}/2-1;k_{21}/2;\mu'\mu/2\right)\right.\right.\\
&\quad-\ {}_1F_1\left(k_{21}/2-2;k_{21}/2;\mu'\mu/2\right)\}]+\\
&\quad\left[\rho^{-2}e^{-\left(\mu'\mu/2\right)}\left(k_{21}-2\right)^{-1}\ {}_1F_1\left(k_{21}/2-1;k_{21}/2;\mu'\mu/2\right)\right].
\end{aligned}
\tag{85}
$$

Finally, noting that ${}_1F_1\left(\alpha;\alpha;z\right)=e^z$, we can simplify the expression above by writing

$$
\begin{aligned}
A &= -e^{-\left(\mu'\mu/2\right)}\left(k_{21}-2\right)^{-1}\ {}_1F_1\left(k_{21}/2-2;k_{21}/2;\mu'\mu/2\right)+e^{-\left(\mu'\mu/2\right)}\ {}_1F_1\left(k_{21}/2-2;k_{21}/2;\mu'\mu/2\right)\\
&\quad+\rho^{-2}e^{-\left(\mu'\mu/2\right)}\left(k_{21}-2\right)^{-1}\ {}_1F_1\left(k_{21}/2-1;k_{21}/2;\mu'\mu/2\right)\\
&= e^{-\left(\mu'\mu/2\right)}\left[\left(k_{21}-3\right)/\left(k_{21}-2\right)\right]\ {}_1F_1\left(k_{21}/2-2;k_{21}/2;\mu'\mu/2\right)+\\
&\quad\rho^{-2}e^{-\left(\mu'\mu/2\right)}\left(k_{21}-2\right)^{-1}\ {}_1F_1\left(k_{21}/2-1;k_{21}/2;\mu'\mu/2\right).
\end{aligned}
\tag{86}
$$

To show (b), first assume that $k_{21}>4$, so that $\frac{\mu'\mu}{2}>0$, $\frac{k_{21}}{2}-1>0$, $\frac{k_{21}}{2}>0$, and $\frac{k_{21}}{2}-2>0$. It follows that we can apply Lemma A2 to each of the confluent hypergeometric functions ${}_1F_1(\cdot;\cdot;\cdot)$ which appear in

(6) to obtain

$$\begin{aligned}
m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\left[\left(\frac{k_{21}-3}{k_{21}-2}\right)\frac{\Gamma(k_{21}/2)}{\Gamma(k_{21}/2-2)}e^{\frac{\mu'\mu}{2}}\left(\frac{\mu'\mu}{2}\right)^{-1}\left(1+O((\mu'\mu)^{-1})\right)\right. \\
&\quad \left.+\rho^{-2}\left(k_{21}-2\right)^{-1}\left[\Gamma(k_{21}/2)/\Gamma(k_{21}/2-1)\right]e^{(\mu'\mu/2)}\left(\mu'\mu/2\right)^{-1}\left(1+O((\mu'\mu)^{-1})\right)\right] \\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left[\left(\left(k_{21}-3\right)/\left(k_{21}-2\right)\right)\left(\left(k_{21}-2\right)/2\right)\left(\left(k_{21}-4\right)/2\right)\left(\mu'\mu/2\right)^{-1}\right. \\
&\quad \left.\times\left(1+O((\mu'\mu)^{-1})\right)+\rho^{-2}2^{-1}\left(\mu'\mu/2\right)^{-1}\left(1+O((\mu'\mu)^{-1})\right)\right] \\
&= O((\mu'\mu)^{-1}). \qquad (87)
\end{aligned}$$

Next, assume that $k_{21}=4$, and observe that, in this case, $e^{-\frac{\mu'\mu}{2}}\left[(k_{21}-3)/(k_{21}-2)\right]{}_1F_1(k_{21}/2-2;k_{21}/2;\mu'\mu/2) = e^{-\frac{\mu'\mu}{2}}(1/2)\ {}_1F_1(0;2;\mu'\mu/2) = e^{-\frac{\mu'\mu}{2}}(1/2) = O\left(e^{-\frac{\mu'\mu}{2}}\right)$. It follows that
$m_{\widehat{\beta}_{IV}}(\mu'\mu,4) = e^{-\frac{\mu'\mu}{2}}(1/2){}_1F_1(0;2;\mu'\mu/2)+\rho^{-2}e^{-\frac{\mu'\mu}{2}}(1/2){}_1F_1(1;2;\mu'\mu/2) = O\left(e^{-\frac{\mu'\mu}{2}}\right)$
$+O\left((\mu'\mu)^{-1}\right) = O\left((\mu'\mu)^{-1}\right)$.

To show (c), note that $\lim_{k_{21}\to\infty}\frac{\left(\frac{1}{2}k_{21}-\left(\frac{1}{2}k_{21}-1\right)\right)\left(\frac{\mu'\mu}{2}\right)}{\left(\frac{1}{2}k_{21}\right)} = \lim_{k_2\to\infty}\left(\frac{\mu'\mu}{k_{21}}\right) = 0$ and $\lim_{k_{21}\to\infty}\frac{\left(\frac{1}{2}k_{21}-\left(\frac{1}{2}k_{21}-2\right)\right)\left(\frac{\mu'\mu}{2}\right)}{\left(\frac{1}{2}k_{21}\right)}$ $= \lim_{k_2\to\infty}\left(\frac{2\mu'\mu}{k_{21}}\right) = 0$. Hence, each of the ${}_1F_1(\cdot;\cdot;\cdot)$ functions appearing in expression (6) satisfies the conditions of Lemma A3 so we may apply this Lemma to obtain $m_{\widehat{\beta}_{IV}}(\mu'\mu,k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}e^{\frac{\mu'\mu}{2}}$ $\left[\left(\left(k_{21}-3\right)/\left(k_{21}-2\right)\right)\left(1+O(k_{21}^{-1})\right)+\rho^{-2}\left(k_{21}-2\right)^{-1}\left(1+O(k_{21}^{-1})\right)\right] = \sigma_{uu}\sigma_{vv}^{-1}\rho^2\times\left[1+O(k_{21}^{-1})\right]\to\sigma_{uu}\sigma_{vv}^{-1}\rho^2$ as $k_{21}\to\infty$.

To show part (d), it suffices to show that $\frac{\partial m_{\widehat{\beta}_{2SLS}}(\mu'\mu,k_{21})}{\partial(\mu'\mu)} < 0$, for all fixed integer $k_{21}\geq 4$. To proceed, write the MSE formula in its infinite series representation as given by (8):

$$\begin{aligned}
m_{\widehat{\beta}_{IV}}(\mu'\mu,k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\left[\left(\frac{1}{\rho^2}\right)\left(\frac{1}{k_{21}-2}\right)\sum_{j=0}^{\infty}\frac{\left(k_{21}/2-1\right)_j}{\left(k_{21}/2\right)_j}\frac{\left(\mu'\mu/2\right)^j}{j!}\right. \\
&\quad \left.+\left(\frac{k_{21}-3}{k_{21}-2}\right)\sum_{j=0}^{\infty}\frac{\left(k_{21}/2-2\right)_j}{\left(k_{21}/2\right)_j}\frac{\left(\mu'\mu/2\right)^j}{j!}\right] \\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\left[\left(\frac{1}{\rho^2}\right)\left(\frac{1}{k_{21}-2}\right)\left(1+\sum_{j=1}^{\infty}\frac{\left(k_{21}/2-1\right)}{\left(k_{21}/2+j-1\right)}\frac{\left(\mu'\mu/2\right)^j}{j!}\right)+\right. \\
&\quad \left.\left(\frac{k_{21}-3}{k_{21}-2}\right)\left(1+\sum_{j=1}^{\infty}\frac{\left(k_{21}/2-1\right)\left(k_{21}/2-2\right)}{\left(k_{21}/2+j-1\right)\left(k_{21}/2+j-2\right)}\frac{\left(\mu'\mu/2\right)^j}{j!}\right)\right]. \qquad (88)
\end{aligned}$$

Now, differentiating (88) term-by-term, we obtain

$$
\begin{aligned}
\frac{\partial m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21})}{\partial(\mu'\mu)} &= \left(\frac{1}{2}\right)\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\left\{\left[\frac{1}{\rho^2}\left(\frac{1}{k_{21}-2}\right)\sum_{j=1}^{\infty}\frac{(k_{21}/2-1)}{(k_{21}/2+j-1)}\frac{j\left(\mu'\mu/2\right)^{j-1}}{j!}+\right.\right. \\
&\quad \left.\left(\frac{k_{21}-3}{k_{21}-2}\right)\left(\sum_{j=1}^{\infty}\frac{(k_{21}/2-1)(k_{21}/2-2)}{(k_{21}/2+j-1)(k_{21}/2+j-2)}\frac{j\left(\mu'\mu/2\right)^{j-1}}{j!}\right)\right] \\
&\quad -\left[\left(\frac{1}{\rho^2}\right)\left(\frac{1}{k_{21}-2}\right)\left(1+\sum_{j=1}^{\infty}\frac{(k_{21}/2-1)}{(k_{21}/2+j-1)}\frac{\left(\mu'\mu/2\right)^{j}}{j!}\right)\right. \\
&\quad \left.\left.+\left(\frac{k_{21}-3}{k_{21}-2}\right)\left(1+\sum_{j=1}^{\infty}\frac{(k_{21}/2-1)(k_{21}/2-2)}{(k_{21}/2+j-1)(k_{21}/2+j-2)}\frac{\left(\mu'\mu/2\right)^{j}}{j!}\right)\right]\right\} \\
&= -0.5\sigma_{uu}\sigma_{vv}^{-1}e^{-\left(\mu'\mu/2\right)}\left(k_{21}-2\right)^{-1}\times \\
&\quad \left[\frac{2}{k_{21}}+\sum_{j=1}^{\infty}\frac{(k_{21}/2-1)}{(k_{21}/2+j)(k_{21}/2+j-1)}\frac{\left(\mu'\mu/2\right)^{j}}{j!}\right]-\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\left(\frac{k_{21}-3}{k_{21}-2}\right) \\
&\quad \times\left[\frac{2}{k_{21}}+\sum_{j=1}^{\infty}\frac{(k_{21}/2-1)(k_{21}/2-2)}{(k_{21}/2+j)(k_{21}/2+j-1)(k_{21}/2+j-2)}\frac{\left(\mu'\mu/2\right)^{j}}{j!}\right] \\
&< 0 \text{ for } k_{21}\geq 4, \qquad (89)
\end{aligned}
$$

where interchanging the operations of differentiations and summation is justified by the absolute and uniform convergence of the infinite series (88) and (89).

Proof of Theorem 4.1: To show part (a), note that direct application of part (a) of Lemma A7 to the bias expression (5) yields

$$
\begin{aligned}
b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left\{\left(1+\tau^2\right)^{-1}-k_{21}^{-1}\left(1+\tau^2\right)^{-1}\left[2-4\left(1+\tau^2\right)^{-1}+2\left(1+\tau^2\right)^{-2}\right]\right. \\
&\quad -k_{21}^{-2}\left(1+\tau^2\right)^{-2}\left[8-28\left(1+\tau^2\right)^{-1}+32\left(1+\tau^2\right)^{-2}-12\left(1+\tau^2\right)^{-3}\right] \\
&\quad \left.-R^*(k_{21})k_{21}^{-1}\left(1+\tau^2\right)^{-2}+O\left(k_{21}^{-3}\right)\right\} \\
&= \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left\{\left(1+\tau^2\right)^{-1}-k_{21}^{-1}\left(1+\tau^2\right)^{-1}\left[2-4\left(1+\tau^2\right)^{-1}+2\left(1+\tau^2\right)^{-2}\right]\right\} \\
&\quad +O\left(k_{21}^{-2}\right) \\
&= \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left\{\left(1+\tau^2\right)^{-1}-2k_{21}^{-1}\left(1+\tau^2\right)^{-1}\left(\tau^2/\left(1+\tau^2\right)\right)^2\right\}+O\left(k_{21}^{-2}\right). \qquad (90)
\end{aligned}
$$

To show part (b), we first rewrite expression (6) as follows

$$
\begin{aligned}
m_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left[\rho^{-2}\left(k_{21}-2\right)^{-1}\ {}_1F_1\left(k_{21}/2-1;\ k_{21}/2;\ \mu'\mu(\tau^2,k_{21})/2\right)e^{-\mu'\mu(\tau^2,k_{21})/2}\right. \\
&\quad +\left(\frac{k_{21}-3}{k_{21}-2}\right)\left(\frac{k_{21}-2}{2}\right)\ {}_1F_1\left(\frac{k_{21}}{2}-2;\frac{k_{21}}{2}-1;\frac{\mu'\mu(\tau^2,k_{21})}{2}\right)e^{-\frac{\mu'\mu(\tau^2,k_{21})}{2}} \\
&\quad \left.-\left(\frac{k_{21}-3}{k_{21}-2}\right)\left(\frac{k_{21}-4}{2}\right)\ {}_1F_1\left(\frac{k_{21}}{2}-1;\frac{k_{21}}{2};\frac{\mu'\mu(\tau^2,k_{21})}{2}\right)e^{-\frac{\mu'\mu(\tau^2,k_{21})}{2}}\right], \qquad (91)
\end{aligned}
$$

where we have made use of the identity $(\gamma-\alpha-1)\ {}_1F_1(\alpha;\gamma;z) = (\gamma-1)\ {}_1F_1(\alpha;\gamma-1;z) - \alpha\ {}_1F_1(\alpha+1;\gamma;z)$ in rewriting expression (6). (See Lebedev (1972), pp. 262, for more details on this and identities involving confluent hypergeometric functions.) Applying the results of Lemma A7 to the confluent hypergeometric functions in expression (91) above, we obtain

$$\begin{aligned}
m_{\widehat{\beta}_{IV}}(\tau^2,k_{21}) \quad = \quad & \sigma_{uu}\sigma_{vv}^{-1}\rho^2\Big[\rho^{-2}k_{21}^{-1}(1-2/k_{21})^{-1}\Big\{\left(1+\tau^2\right)^{-1} - k_{21}^{-1}\left(1+\tau^2\right)^{-1}\Big(2-4\left(1+\tau^2\right)^{-1} \\
& +2\left(1+\tau^2\right)^{-2}\Big) - k_{21}^{-2}\left(1+\tau^2\right)^{-2}\Big(8-28\left(1+\tau^2\right)^{-1}+32\left(1+\tau^2\right)^{-2} \\
& -12\left(1+\tau^2\right)^{-3}\Big) - R^*(k_{21})k_{21}^{-1}\left(1+\tau^2\right)^{-2} + O\left(k_{21}^{-3}\right)\Big\} + (1-2/k_{21})^{-1}\times \\
& (1-3/k_{21})(k_{21}/2-1)\Big\{\left(1+\tau^2\right)^{-1} - k_{21}^{-1}\left(1+\tau^2\right)^{-1}\Big(4-6\left(1+\tau^2\right)^{-1}+ \\
& 2\left(1+\tau^2\right)^{-2}\Big) - k_{21}^{-2}\left(1+\tau^2\right)^{-2}\Big(24-56\left(1+\tau^2\right)^{-1}+44\left(1+\tau^2\right)^{-2} \\
& -12\left(1+\tau^2\right)^{-3}\Big) - R^*(k_{21})k_{21}^{-1}\left(1+\tau^2\right)^{-2} + O\left(k_{21}^{-3}\right)\Big\} - \\
& (1-3/k_{21})(1-2/k_{21})^{-1}(k_{21}/2-2)\Big\{\left(1+\tau^2\right)^{-1} - k_{21}^{-1}\left(1+\tau^2\right)^{-1}(2 \\
& -4\left(1+\tau^2\right)^{-1}+2\left(1+\tau^2\right)^{-2}\Big) - k_{21}^{-2}\left(1+\tau^2\right)^{-2}\Big(8-28\left(1+\tau^2\right)^{-1}+ \\
& 32\left(1+\tau^2\right)^{-2}-12\left(1+\tau^2\right)^{-3}\Big) - R^*(k_{21})k_{21}^{-1}\left(1+\tau^2\right)^{-2} + O\left(k_{21}^{-3}\right)\Big\} \qquad (92)
\end{aligned}$$

Expanding $(1-2/k_{21})^{-1}$ in the binomial series series $(1-2/k_{21})^{-1} = 1+2/k_{21}+4/k_{21}^2+O\left(k_{21}^{-3}\right)$; and, after some tedious but straightforward calculations, it can be shown that

$$\begin{aligned}
m_{\widehat{\beta}_{IV}}(\tau^2,k_{21}) \quad = \quad & \sigma_{uu}\sigma_{vv}^{-1}\rho^2\Big\{\left(1+\tau^2\right)^{-2} + \rho^{-2}k_{21}^{-1}\left(1+\tau^2\right)^{-1} - k_{21}^{-1}\left(1+\tau^2\right)^{-2}[7 \\
& -12\left(1+\tau^2\right)^{-1}+6\left(1+\tau^2\right)^{-2}\Big] + O(k_{21}^{-2})\Big\} \\
= \quad & \sigma_{uu}\sigma_{vv}^{-1}\rho^2\Big\{\left(1+\tau^2\right)^{-2} + \left(\left(1-\rho^2\right)/\rho^2\right)k_{21}^{-1}\left(1+\tau^2\right)^{-1} + k_{21}^{-1}\left(1+\tau^2\right)^{-1}\times \\
& \Big[1-7\left(1+\tau^2\right)^{-1}+12\left(1+\tau^2\right)^{-2}-6\left(1+\tau^2\right)^{-3}\Big]\Big\} + O(k_{21}^{-2}). \qquad (93)
\end{aligned}$$

Proof of Theorem 4.3: To prove (a), write

$$\begin{aligned}
\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2,k_{21}) \quad = \quad & \sigma_{uu}\sigma_{vv}^{-1}\rho^2\Big\{\left(1+\tau^2\right)^{-2} + \left(\left(1-\rho^2\right)/\rho^2\right)k_{21}^{-1}\left(1+\tau^2\right)^{-1} + \\
& k_{21}^{-1}\left(1+\tau^2\right)^{-1}\Big[1-7\left(1+\tau^2\right)^{-1}+12\left(1+\tau^2\right)^{-2}-6\left(1+\tau^2\right)^{-3}\Big]\Big\} \\
= \quad & \sigma_{uu}\sigma_{vv}^{-1}\rho^2 k_{21}^{-1}\left(1+\tau^2\right)^{-1}\Big[\rho^{-2}-(7-k_{21})\left(1+\tau^2\right)^{-1} \\
& +12\left(1+\tau^2\right)^{-2}-6\left(1+\tau^2\right)^{-3}\Big] \qquad (94)
\end{aligned}$$

.

Now, let $x = \left(1+\tau^2\right)^{-1}$ and define the polynomial $\varphi(x) = \rho^{-2}-(7-k_{21})x+12x^2-6x^3$. From expression (94) above, it is apparent that to show that $\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2,k_{21}) \geq 0$ for $\tau^2 \in [0,\infty)$ and $k_{21} \geq 4$, it suffices to show that $\varphi(x) \geq 0$ for $x \in [0,1]$ and $k_{21} \geq 4$. Note also that, for $k_{21} \geq 4$ and $x \in [0,1]$, $\varphi(x) \geq 1-3x+12x^2-6x^3 = \pi(x)$ (*say*); so that we need only to show that $\pi(x) \geq 0$ for $x \in [0,1]$. We will, in fact, show that $\pi(x) > 0$ for

$x \in [0,1]$. To proceed, observe first that $\pi(0) = 1$ and $\pi(1) = 4$, so that $\pi(.)$ is positive at the end points of the interval. Next, to show that $\pi(x)$ is positive in the open interval $(0,1)$, we note that $\pi'(x) = -3 + 24x - 18x^2$, so that $\pi'(x)$ is a continuously differentiable (quadratic) function which is 0 at $x = 2/3 - \left(\sqrt{10}\right)/6 \approx .14$, negative for $x \in \left[0, 2/3 - \left(\sqrt{10}\right)/6\right)$, and positive for $x \in \left(2/3 - \left(\sqrt{10}\right)/6, 1\right]$. Moreover, observe that $\pi\left(2/3 - \left(\sqrt{10}\right)/6\right) = 1 - 3\left(2/3 - \left(\sqrt{10}\right)/6\right) + 12\left(2/3 - \left(\sqrt{10}\right)/6\right)^2 - 6\left(2/3 - \left(\sqrt{10}\right)/6\right)^3$ $= 23/9 - (5/9)\sqrt{10} \approx .80 > 0$. It follows that $\pi(x)$ is monotonically decreasing for $x \in \left[0, 2/3 - \left(\sqrt{10}\right)/6\right)$ reaching a local minimal value of $23/9 - (5/9)\sqrt{10}$ at the point $x = 2/3 - \left(\sqrt{10}\right)/6$. Furthermore, $\pi(x)$ is monotonically increasing for $x \in \left(2/3 - \left(\sqrt{10}\right)/6, 1\right]$ reaching a value of 4 at $x = 1$. It, thus, follows immediately that $\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) \geq 0$ for $\tau^2 \in [0,\infty)$ and $k_{21} \geq 4$, and note that $\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) = 0$ only if $\sigma_{uv} = 0$.

The proof for part (b) follows in a similar manner as that given by part (a). We, thus, omit the details.

Proof of Theorem 4.4: To show part (a), we write the infinite series representation of expression (29) as

$$\begin{aligned} b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \rho e^{-\frac{\tau^2 k_{21}}{2}} \left[\sum_{j=0}^{\infty} \left(\frac{k_{21} - 2}{k_{21} + 2j - 2} \right) \frac{\left(0.5\tau^2 k_{21}\right)^j}{j!} \right] \\ &= \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \rho f(\tau^2, k_{21}) \ (say). \end{aligned} \tag{95}$$

Note that $f(\tau^2 = 0, k_{21}) = 1$, and in a manner similar to the proof of part (b) of Theorem 3.1, we can show that $\lim_{\tau^2 \to \infty} f(\tau^2, k_{21}) = 0$ for k_{21} fixed. Moreover, taking partial derivative of $f(\tau^2, k_{21})$ with respect to τ^2, we obtain

$$\begin{aligned} \frac{\partial f(\tau^2, k_{21})}{\partial \tau^2} &= -\left(\frac{k_{21}}{2}\right) e^{-\frac{\tau^2 k_{21}}{2}} \sum_{j=0}^{\infty} \left(\frac{k_{21} - 2}{k_{21} + 2j - 2} \right) \frac{\left(0.5\tau^2 k_{21}\right)^j}{j!} \\ &\quad + \left(\frac{k_{21}}{2}\right) e^{-\frac{\tau^2 k_{21}}{2}} \sum_{j=0}^{\infty} \left(\frac{k_{21} - 2}{k_{21} + 2j - 2} \right) \frac{j\left(0.5\tau^2 k_{21}\right)^{j-1}}{j!} \\ &= \left(\frac{k_{21}}{2}\right) e^{-\frac{\tau^2 k_{21}}{2}} \sum_{j=0}^{\infty} \frac{(k_{21} - 2)\left[k_{21} + 2j - 2 - (k_{21} + 2j)\right]}{(k_{21} + 2j)(k_{21} + 2j - 2)} \frac{\left(0.5\tau^2 k_{21}\right)^j}{j!} \\ &= -k_{21} e^{-\frac{\tau^2 k_{21}}{2}} \sum_{j=0}^{\infty} \frac{(k_{21} - 2)}{(k_{21} + 2j)(k_{21} + 2j - 2)} \frac{\left(0.5\tau^2 k_{21}\right)^j}{j!} < 0, \end{aligned} \tag{96}$$

where the term-by-term differentiation is justified by the absolute and uniform convergence of the series involved. It follows that $0 \leq f(\tau^2, k_{21}) \leq 1$, and $f(\tau^2, k_{21})$ is a monotonically decreasing function in the range $\tau^2 \in [0,\infty)$ for k_{21} fixed. Now, note that the sign of $b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ is determined by the sign of ρ. Hence, from expression (95), we see that, for $\rho > 0$, $\frac{\partial b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})}{\partial \tau^2} = \rho \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \frac{\partial f(\tau^2, k_{21})}{\partial \tau^2} < 0$ and, for $\rho < 0$, $\frac{\partial b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})}{\partial \tau^2} = \rho \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \frac{\partial f(\tau^2, k_{21})}{\partial \tau^2} > 0$; so that $\left| b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) \right|$ is a monotonically decreasing function of τ^2 for $\rho \neq 0$ (or equivalently for $\sigma_{uv} \neq 0$), which establishes the desired result.

To show part (b), we note that, making use of the infinite series representation of expression (76), reparameterized in terms of τ^2 and k_{21}, we can rewrite the bias formula (29) as:

$$
\begin{aligned}
b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) &= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1 - e^{-\frac{\tau^2 k_{21}}{2}}\sum_{j=0}^{\infty}\frac{\left(0.5\tau^2 k_{21}\right)^{j+1}}{j!}\frac{1}{(k_{21}/2+j)}\right] \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho g(\tau^2, k_{21}),\ (say). \qquad (97)
\end{aligned}
$$

Now, taking the partial derivative of $g(\tau^2, k_{21})$ in expression (97) with respect to k_{21}, we obtain

$$
\begin{aligned}
\frac{\partial g(\tau^2, k_{21})}{\partial k_{21}} &= e^{-\frac{\tau^2 k_{21}}{2}}\left[\left(\frac{\tau^2}{2}\right)\sum_{j=0}^{\infty}\frac{\left(0.5\tau^2 k_{21}\right)^{j+1}}{j!}\frac{1}{(k_{21}/2+j)} + \left(\frac{1}{2}\right)\sum_{j=0}^{\infty}\frac{\left(0.5\tau^2 k_{21}\right)^{j+1}}{j!}\frac{1}{(k_{21}/2+j)^2}\right. \\
&\quad \left. - \left(\frac{\tau^2}{2}\right)\sum_{j=0}^{\infty}\frac{(j+1)\left(0.5\tau^2 k_{21}\right)^{j}}{j!}\frac{1}{(k_{21}/2+j)}\right], \qquad (98)
\end{aligned}
$$

where term-by-term differentiation above is justified by the absolute and uniform convergence of the series involved. Now, take $i = j+1$, we see that

$$
\begin{aligned}
\frac{\partial g(\tau^2, k_{21})}{\partial k_{21}} &= e^{-\frac{\tau^2 k_{21}}{2}}\left[\left(\frac{\tau^2}{2}\right)\sum_{i=0}^{\infty}\frac{i\left(0.5\tau^2 k_{21}\right)^{i}}{i!}\frac{(k_{21}/2+i)}{(k_{21}/2+i)(k_{21}/2+i-1)}\right. \\
&\quad - \left(\frac{\tau^2}{2}\right)\sum_{j=0}^{\infty}\frac{j\left(0.5\tau^2 k_{21}\right)^{j}}{j!}\frac{(k_{21}/2+j-1)}{(k_{21}/2+j)(k_{21}/2+j-1)} \\
&\quad - \left(\frac{\tau^2}{2}\right)\sum_{j=0}^{\infty}\frac{\left(0.5\tau^2 k_{21}\right)^{j}}{j!}\frac{(k_{21}/2+j-1)}{(k_{21}/2+j)(k_{21}/2+j-1)} \\
&\quad \left. + \left(\frac{1}{2}\right)\sum_{j=0}^{\infty}\frac{\left(0.5\tau^2 k_{21}\right)^{j+1}}{j!}\frac{1}{(k_{21}/2+j)^2}\right] \\
&= e^{-\frac{\tau^2 k_{21}}{2}}\left[-\left(\frac{\tau^2}{2}\right)\sum_{j=0}^{\infty}\frac{\left(0.5\tau^2 k_{21}\right)^{j}}{j!}\frac{(k_{21}/2-1)}{(k_{21}/2+j)(k_{21}/2+j-1)}\right. \\
&\quad \left. + \left(\frac{1}{2}\right)\sum_{j=0}^{\infty}\frac{\left(0.5\tau^2 k_{21}\right)^{j+1}}{j!}\frac{1}{(k_{21}/2+j)^2}\right] \\
&= \left(\frac{1}{2}\right)e^{-\frac{\tau^2 k_{21}}{2}}\left[\sum_{j=0}^{\infty}\frac{\left(0.5\tau^2 k_{21}\right)^{j}}{j!}\frac{\tau^2 j}{(k_{21}/2+j)^2(k_{21}/2+j-1)}\right] > 0. \qquad (99)
\end{aligned}
$$

Moreover, similar to the proof of part (a) above, we have that the sign of $b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ is determined by the sign of ρ. Hence, $\frac{\partial b_{\widehat{\beta}_{IV}}(\tau^2,k_{21})}{\partial k_{21}} = \rho\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\frac{\partial g(\tau^2,k_{21})}{\partial k_{21}} > 0$ for $\rho > 0$ and $\frac{\partial b_{\widehat{\beta}_{IV}}(\tau^2,k_{21})}{\partial k_{21}} = \rho\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\frac{\partial g(\tau^2,k_{21})}{\partial k_{21}} < 0$ for $\rho < 0$, and $|b_{\widehat{\beta}_{IV}}(\tau^2, k_{21})|$ is a monotonically increasing function in k_{21} for $\rho \neq 0$, and part (b) is, thus, proved.

Finally, to show part (c), we note that, analogous to expression (88) above, the MSE formula can be written as

$$m_{\widehat{\beta}_{IV}}(\tau^2,k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\tau^2 k_{21}}{2}}\left[\left(\frac{1}{\rho^2}\right)\left(\frac{1}{k_{21}-2}\right)\left(1+\sum_{j=1}^{\infty}\frac{(k_{21}/2-1)}{(k_{21}/2+j-1)}\frac{\left(0.5\tau^2 k_{21}\right)^j}{j!}\right)+\right.$$
$$\left.\left(\frac{k_{21}-3}{k_{21}-2}\right)\left(1+\sum_{j=1}^{\infty}\frac{(k_{21}/2-1)(k_{21}/2-2)}{(k_{21}/2+j-1)(k_{21}/2+j-2)}\frac{\left(0.5\tau^2 k_{21}\right)^j}{j!}\right)\right]. \tag{100}$$

Next, differentiating expression (100) term-by-term with respect to τ^2, we obtain

$$\begin{aligned}
\frac{\partial m_{\widehat{\beta}_{IV}}(\tau^2,k_{21})}{\partial\tau^2} &= -\frac{k_{21}}{2}\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\tau^2 k_{21}}{2}}\left[\left(\frac{1}{\rho^2}\right)\left(\frac{1}{k_{21}-2}\right)\left(1+\sum_{j=1}^{\infty}\frac{(k_{21}-2)}{(k_{21}+2j-2)}\frac{\left(0.5\tau^2 k_{21}\right)^j}{j!}\right)\right. \\
&\quad \left.+\left(\frac{k_{21}-3}{k_{21}-2}\right)\left(1+\sum_{j=1}^{\infty}\frac{(k_{21}-2)(k_{21}-4)}{(k_{21}+2j-2)(k_{21}+2j-4)}\frac{\left(0.5\tau^2 k_{21}\right)^j}{j!}\right)\right]+ \\
&\quad \frac{k_{21}}{2}\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\tau^2 k_{21}}{2}}\left[\left(\frac{1}{\rho^2}\right)\left(\frac{1}{k_{21}-2}\right)\left(\sum_{j=1}^{\infty}\frac{(k_{21}-2)}{(k_{21}+2j-2)}\frac{j\left(0.5\tau^2 k_{21}\right)^{j-1}}{j!}\right)\right. \\
&\quad \left.+\left(\frac{k_{21}-3}{k_{21}-2}\right)\left(\sum_{j=1}^{\infty}\frac{(k_{21}-2)(k_{21}-4)}{(k_{21}+2j-2)(k_{21}+2j-4)}\frac{j\left(0.5\tau^2 k_{21}\right)^{j-1}}{j!}\right)\right] \\
&= -\frac{k_{21}}{2}\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\tau^2 k_{21}}{2}}\times \\
&\quad \left[\left(\frac{1}{\rho^2}\right)\left(\frac{1}{k_{21}-2}\right)\left(\sum_{j=0}^{\infty}\frac{2(k_{21}-2)}{(k_{21}+2j)(k_{21}+2j-2)}\frac{\left(0.5\tau^2 k_{21}\right)^j}{j!}\right)\right. \\
&\quad \left.+\left(\frac{k_{21}-3}{k_{21}-2}\right)\left(\sum_{j=0}^{\infty}\frac{4(k_{21}-2)(k_{21}-4)}{(k_{21}+2j)(k_{21}+2j-2)(k_{21}+2j-4)}\frac{\left(0.5\tau^2 k_{21}\right)^j}{j!}\right)\right] \\
&< 0,
\end{aligned} \tag{101}$$

where again the interchanging of differentiation and summation is justified by the absolute and uniform convergence of the infinite series (100) and (101) for $k_{21}\geq 4$.

Proof of Lemma 5.1: To show part (a), 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 A8 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 (a) follow as an immediate consequence.

To show part (b), note that it follows from part (d) of Lemma A8, part (a) of this Lemma, and the continuous mapping theorem that as $T\to\infty$, $W_{k_{21},T}=\frac{\left(\frac{y_2' M_X Z_1}{\sqrt{T}}\right)\left(\frac{Z_1' M_X Z_1}{T}\right)^{-1}\left(\frac{Z_1' M_X y_2}{\sqrt{T}}\right)/k_{21}}{\widehat{\sigma}_{vv,1}}\Longrightarrow\frac{(\mu+Z_{v,1})'(\mu+Z_{v,1})}{k_{21}}=\frac{v_1(\mu'\mu,k_{21})}{k_{21}}$. It then follows directly from part (a) of Lemma A9 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 (c), write $\widehat{\sigma}_{uv,1}=\frac{\left(y_1-y_2\widehat{\beta}_{IV}\right)' M_{(Z,X)} y_2}{T}\left(\frac{W_{k_{21},T}}{W_{k_{21},T}-1}\right)=\left[\frac{u' M_{(Z,X)} y_2}{T}-\left(\widehat{\beta}_{IV}-\beta_0\right)\right.$

$\frac{y_2' M_{(Z,X)} y_2}{T}\Big] \left(\frac{W_{k_{21},T}}{W_{k_{21},T}-1}\right)$ 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 A8, 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 A9, 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.

To show part (d), we 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 A8; the proofs of parts (a), (b), and (c) of this Lemma as given above; and the continuous mapping theorem that as $T \to \infty$, $\widehat{\sigma}_{uu,1} \Longrightarrow \sigma_{uu} - 2\sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \sigma_{uv} \left(\frac{v_1(\mu'\mu, k_{21})}{k_{21}}\right)^{-1} \left(\frac{v_2(\mu'\mu, k_{21})}{k_{21}}\right) + \sigma_{uu}\left(\frac{v_1(\mu'\mu, k_{21})}{k_{21}}\right)^{-2} \left(\frac{v_2(\mu'\mu, k_{21})}{k_{21}}\right)^2 + 2\frac{\mathcal{A}_{k_{21},\mu'\mu}^2}{\sigma_{vv}} \left(\frac{v_1(\mu'\mu, k_{21})}{k_{21}}\right)^{-1} - \frac{\mathcal{A}_{k_{21},\mu'\mu}^2}{\sigma_{vv}} \left(\frac{v_1(\mu'\mu, k_{21})}{k_{21}}\right)^{-2} = \mathcal{B}_{k_{21},\mu'\mu}$ (say). Similarly, $\widehat{\sigma}_{uu,2} \Longrightarrow \mathcal{B}_{k_{21},\mu'\mu}$ as $T \to \infty$. Moreover, applying Lemma A9 and part (c) 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 results.

Proof of Theorem 5.2: For each part of this theorem, we will only prove the convergence result for the estimator with subscript $i = 1$ since the proofs for the estimators with subscript $i = 2$ follow in a like manner. First, to show (a) for the case $i = 1$, 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 the proofs of parts (a), (b), and (c) of Lemma 5.1 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 A9 and part (c) of Lemma 5.1, we deduce 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 (b) for the case $i = 1$, write $\widetilde{BIAS}_1 = \widehat{BIAS}_1 - \left(\frac{2}{k_{21}}\right)\left[\left(\frac{\widehat{\sigma}_{uv,1}}{\widehat{\sigma}_{vv,1}}\right)\left(\frac{1}{W_{k_{21},T}}\right)\left(\frac{W_{k_{21},T}-1}{W_{k_{21},T}}\right)^2\right]$. Again, note that from the proofs of parts (a), (b), and (c) of Lemma 5.1 and the continuous mapping theorem, we see that as $T \to \infty$, $\widetilde{BIAS}_1 \Longrightarrow \mathcal{C}_{k_{21},\mu'\mu} - \frac{2}{k_{21}}\left[\frac{\mathcal{A}_{k_{21},\mu'\mu}}{\sigma_{vv}}\left(\frac{v_1(\mu'\mu, k_{21})}{k_{21}}\right)^{-1}\left(\frac{v_1(\mu'\mu, k_{21})}{k_{21}}\right)^{-2}\left(\frac{v_1(\mu'\mu, k_{21})}{k_{21}} - 1\right)^2\right]$
$= \mathcal{E}_{k_{21},\mu'\mu}$ (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 A9, part (c) of Lemma 5.1, and part (a) of this theorem; we deduce 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 (c) for the case $i = 1$, write $\widehat{MSE}_1 = \left(\widehat{BIAS}_1\right)^2$. It follows immediately from the proof of part (a) above and the continuous mapping theorem that as $T \to \infty$, $\widehat{MSE}_1 \Longrightarrow \mathcal{C}_{k_{21},\mu'\mu}^2$. Moreover, given part (a) of this theorem, as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4, we deduce easily that $\mathcal{C}_{k_{21},\mu'\mu}^2 \xrightarrow{p} \sigma_{uu}\sigma_{vv}\rho^2\left(\frac{1}{1+\tau^2}\right)^2$.

To show part (d) for the case $i = 1$, write $\widetilde{MSE}_1 = \widehat{MSE}_1 + \frac{1}{k_{21}}\left(\frac{\widehat{\sigma}_{uv,1}^2}{\widehat{\sigma}_{vv,1}^2}\right)\left(\frac{1}{W_{k_{21},T}}\right)$
$\left[\left(\frac{\widehat{\sigma}_{uu,1}\widehat{\sigma}_{vv,1} - \widehat{\sigma}_{uv,1}^2}{\widehat{\sigma}_{uv,1}^2}\right) + \left(1 - \frac{7}{W_{k_{21},T}} + \frac{12}{W_{k_{21},T}^2} - \frac{6}{W_{k_{21},T}^3}\right)\right]$. Hence, from the proof of Lemma 5.1, the proof of

part (c) of this theorem given above, and the continous mapping theorem; it is apparent that as $T \to \infty$, $\widetilde{MSE}_1 \Longrightarrow \mathcal{C}^2_{k_{21},\mu'\mu} + \frac{1}{k_{21}} \left(\frac{\mathcal{A}^2_{k_{21},\mu'\mu}}{\sigma^2_{vv}} \right) \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-1}$
$\left[\left(\frac{\sigma_{vv}\mathcal{B}_{k_{21},\mu'\mu} - \mathcal{A}^2_{k_{21},\mu'\mu}}{\mathcal{A}^2_{k_{21},\mu'\mu}} \right) + \left(1 - 7 \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-1} + 12 \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-2} - 6 \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-3} \right) \right] = \mathcal{F}_{k_{21},\mu'\mu}\ (say).$
Futhermore, 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 A9, parts (c) and (d) of Lemma 5.1, and part (c) 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.

To show part (e) for the case $i = 1$, we note that by comparing expression (35) for $\widetilde{MSE}_1$ with expression (36) for $\overline{MSE}_1$, we see that the only difference between these two alternative estimators for the MSE, as explained in Subsection 5.1, is that $\widetilde{MSE}_1$ estimates the quantity $\sigma_{uu}\sigma_{vv} - \sigma_{uv}$ using the consistent estimator $\widehat{\sigma}_{uu,1}\widehat{\sigma}_{vv,1} - \widehat{\sigma}^2_{uv,1}$ whereas $\overline{MSE}_1$ estimates the quantity $g_{11}g_{22} - g^2_{12}$ using the estimator $\widehat{g}_{11}\widehat{g}_{22} - \widehat{g}^2_{12}$. Since it is easy to verify that $g_{11}g_{22} - g^2_{12} = \sigma_{uu}\sigma_{vv} - \sigma_{uv}$, all that is left to show is the consistency of the estimator $\widehat{g}_{11}\widehat{g}_{22} - \widehat{g}^2_{12}$. 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, $\widehat{g}_{11}\widehat{g}_{22} - \widehat{g}^2_{12} \xrightarrow{p} g_{11}g_{22} - g^2_{12} = \sigma_{uu}\sigma_{vv} - \sigma_{uv}$. Since these limits do not depend on k_{21} and $\mu'\mu$, the desired result follows as a direct consequence.

Proof of Lemma 5.4: To begin, we note that since $\widehat{\sigma}_{vv,1} = \frac{y_2' M_{(Z,X)} y_2}{T}$ and $\widehat{\sigma}_{vv,2} = \frac{y_2' M_X y_2}{T}$, parts (a) and (b) of this lemma follows immediately from parts (b) and (a), respectively, of Lemma A10.

To show part (c), write $W_{k_{21},T} = \left[\frac{y_2'(P_{(Z_1,X)} - P_X) y_2}{\widehat{\sigma}_{vv,1}} \right] k_{21}^{-1} = \frac{T}{\widehat{\sigma}_{vv,1} k_{21}} \left[\frac{y_2' M_X Z_1}{T} \left(\frac{Z_1' M_X Z_1}{T} \right)^{-1} \frac{Z_1' M_X y_2}{T} \right]$. Thus, it follows directly from part (c) of Lemma A10, part (a) of this lemma, and the Slutsky's theorem that $W_{k_{21},T} = O_p(T)$.

To show part (d), first consider the estimator $\widehat{\sigma}_{uv,1}$. Note that, given part (b) of Lemma A10, the well-known consistency of $\widehat{\beta}_{IV}$ under Assumption 1*, and the Slutsky's Theorem; we deduce that as $T \to \infty$, $s_{uv,1} = \frac{(y_1 - y_2\widehat{\beta}_{IV})' M_{(Z,X)} y_2}{T} = \frac{u' M_{(Z,X)} y_2}{T} + \left(\beta_0 - \widehat{\beta}_{IV}\right) \frac{y_2' M_{(Z,X)} y_2}{T} \xrightarrow{p} \sigma_{uv}$. It follows immediately from part (c) of this lemma and the Slutsky's theorem that $\widehat{\sigma}_{uv,1} = s_{uv,1} \left(\frac{1}{1 - \frac{1}{W_{k_{21},T}}} \right) \xrightarrow{p} \sigma_{uv}$, as $T \to \infty$. Next, consider the estimator $\widehat{\sigma}_{uv,2}$. Here, write $s_{uv,2} = \frac{(y_1 - y_2\widehat{\beta}_{IV})' M_X y_2}{T} = \frac{u' M_X y_2}{T} + \left(\beta_0 - \widehat{\beta}_{IV}\right) \frac{y_2' M_X y_2}{T}$; and note that part (a) of Lemma A10, the consistency of $\widehat{\beta}_{IV}$ under Assumption 1*, and the Slutsky's Theorem together imply that as $T \to \infty$, $s_{uv,2} \xrightarrow{p} \sigma_{uv}$. Hence, it follows immediately from part (c) of this lemma and the Slutsky's Theorem that as $T \to \infty$, $\widehat{\sigma}_{uv,i} = s_{uv,i} \left(\frac{1}{1 - \frac{1}{W_{k_{21},T}}} \right) \xrightarrow{p} \sigma_{uv}$, for $i = 1, 2$.

To show part (e), note first that both $\widehat{\sigma}_{uu,1}$ and $\widehat{\sigma}_{uu,2}$ depend on s_{uu}. Note further that $s_{uu} = \frac{(y_1 - y_2\widehat{\beta}_{IV})' M_X (y_1 - y_2\widehat{\beta}_{IV})}{T} = \frac{u' M_X u}{T} + 2\left(\beta_0 - \widehat{\beta}_{IV}\right) \frac{y_2' M_X u}{T} + \left(\beta_0 - \widehat{\beta}_{IV}\right)^2 \frac{y_2' M_X y_2}{T} \xrightarrow{p} \sigma_{uu}$ as $T \to \infty$ as a direct consequence of part (a) of Lemma A10, the consistency of $\widehat{\beta}_{IV}$ under Assumption 1*, and the Slutsky's theorem. Next, consider the estimator $\widehat{\sigma}_{uu,1} = s_{uu} + 2\frac{\widehat{\sigma}^2_{uv,1}}{\widehat{\sigma}_{vv,1}} \left(\frac{1}{W_{k_{21},T}} \right) - \frac{\widehat{\sigma}^2_{uv,1}}{\widehat{\sigma}_{vv,1}} \left(\frac{1}{W_{k_{21},T}} \right)^2$. In view of parts (a), (c), and (d) of this lemma and the Slutsky's Theorem, it is apparent that $\widehat{\sigma}_{uu,1} = s_{uu} + Op\left(\frac{1}{T}\right)$, so that $\widehat{\sigma}_{uu,1} \xrightarrow{p} \sigma_{uu}$ as $T \to \infty$. Finally, consider the estimator $\widehat{\sigma}_{uu,2} = s_{uu} + 2\frac{\widehat{\sigma}^2_{uv,2}}{\widehat{\sigma}_{vv,2}} \left(\frac{1}{W_{k_{21},T}} \right) - \frac{\widehat{\sigma}^2_{uv,2}}{\widehat{\sigma}_{vv,2}} \left(\frac{1}{W_{k_{21},T}} \right)^2$. In this case, parts (b), (c), and (d) of this lemma and the Slutsky's Theorem imply that $\widehat{\sigma}_{uu,2} = s_{uu} + Op\left(\frac{1}{T}\right)$, so that $\widehat{\sigma}_{uu,2} \xrightarrow{p} \sigma_{uu}$ as $T \to \infty$.

Proof of Theorem 5.5: The results of parts (a)-(d) follow as a direct consequence of the results of Lemma 5.4 and the Slutsky's theorem. Moreover, to show (e) note that under Assumption 1*, it is well-known that, as $T \to \infty$, $\widehat{g}_{11} = \frac{y_1' M_{(Z,X)} y_1}{T} \xrightarrow{p} g_{11}$, $\widehat{g}_{12} = \frac{y_1' M_{(Z,X)} y_2}{T} \xrightarrow{p} g_{12}$, and $\widehat{g}_{22} = \frac{y_2' M_{(Z,X)} y_2}{T} \xrightarrow{p} g_{22}$. Hence, applying Lemma 5.4 and the Slutsky's Theorem, we deduce that $\overline{MSE}_1 = Op\left(\frac{1}{T}\right)$ and that $\overline{MSE}_2 = Op\left(\frac{1}{T}\right)$.

Proof of Theorem 5.6: We will only prove consistency results for $\widetilde{\beta}_{OLS,1}$, $\widetilde{\beta}_{IV}$, and $\widetilde{\widetilde{\beta}}_{IV,1}$ since the results for $\widetilde{\beta}_{OLS,2}$ and $\widetilde{\widetilde{\beta}}_{IV,2}$ can be shown in a manner similar to those for $\widetilde{\beta}_{OLS,1}$ and $\widetilde{\widetilde{\beta}}_{IV,1}$, respectively. To prove part (a) for the estimator $\widetilde{\beta}_{OLS,1}$, 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 A8**,** the proof of parts (a) and (c) of Lemma 5.1, 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 (c) of Lemma 5.1.

To show part (b), write $\widetilde{\beta}_{IV} = \widehat{\beta}_{IV} - \widehat{BIAS}_1$. Making use of Lemma A1, the proof of part (a) of Theorem 5.2, 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 A9 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 5.2 part (a); 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 (c) for the estimator $\widetilde{\widetilde{\beta}}_{IV,1}$, write $\widetilde{\widetilde{\beta}}_{IV,1} = \widehat{\beta}_{IV} - \widetilde{BIAS}_1$. Now, given Lemma A1, the proof of part (b) of Theorem 5.2, and the continuous mapping theorem; it is apparent that, as $T \to \infty$, $\widetilde{\widetilde{\beta}}_{IV,1} \Longrightarrow \beta_0 + \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \left[\frac{v_2(\mu'\mu, k_{21})}{v_1(\mu'\mu, k_{21})}\right] - \mathcal{E}_{k_{21},\mu'\mu} = \mathcal{N}_{k_{21},\mu'\mu}$ *(say)*. Moreover, applying Lemma A9 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 5.2 part (b), 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 5.7: Note, of course, that under Assumption 1*, the SEM described in Section 2 is fully identified in the usual sense. Hence, it is well-known by standard arguments that $\widehat{\beta}_{OLS} = \beta + (y_2 M_X y_2/T)^{-1} (y_2 M_X u/T) \xrightarrow{p} \beta_0 + \frac{\sigma_{uv}}{\Pi'\Omega\Pi + \sigma_{vv}}$ and $\widehat{\beta}_{IV} = \beta + \left(\frac{y_2 M_X Z_1}{T} \left(\frac{Z_1' M_X Z_1}{T}\right)^{-1} \frac{Z_1' M_X y_2}{T}\right)^{-1} \left(\frac{y_2 M_X Z_1}{T} \left(\frac{Z_1' M_X Z_1}{T}\right)^{-1} \frac{Z_1' M_X u}{T}\right) \xrightarrow{p} \beta_0$ as $T \to \infty$, as can be seen from direct application of parts (a) and (c) of Lemma A10 and the Slutsky's Theorem. Parts (a) and (b) of this theorem then follows as a direct consequence of parts (a), (b), and (d) of Lemma 5.4, the probability limit of $\widehat{\beta}_{OLS}$ given above, and the Slutsky's Theorem. On the other hand, parts (c) and (d) of this Theorem follows as a direct consequence of parts (a)-(d) of Lemma 5.4, the consistency of $\widehat{\beta}_{IV}$ under full identification, and the Slutsky's Theorem. The arguments are standard, and so we omit the details.

References

Angrist, J. D. and A. B. Krueger (1991). "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106, 979-1014.

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.

Billingsley, P. (1968). *Convergence of Probability Measures* (John Wiley, New York).

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.

Chao, J.C. and P.C.B.Phillips (1998). "Posterior Distributions in Limited Information Analysis of the Simultaneous Equations Model Using the Jeffreys Prior," *Journal of Econometrics*, 87, 49-86.

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 (2001). "Pathwise Asymptotics for Bias-Corrected Single-Equation Estimators When the Number of Weak Instruments Is Large," unpublished manuscript, University of Maryland.

De Bruijn, N. G. (1961). *Asymptotic Methods in Analysis* (North-Holland, Amsterdam).

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

Elliott, G. (1998). "The Robustness of Cointegration Methods when Regressors Almost Have Unit Roots," *Econometrica*, 66, pp149-58, 1998.

Hahn, J. (1997). "Optimal Inference with Many Instruments," University of Brown 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.

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.

Judge, G. and M.E. Bock (1978). *The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics* (North Holland, New York).

Kass, R.E., L. Tierney and J.B.Kadane (1990). "The Validity of Posterior Expansions Based on Laplace's Method," In: Geisser, S., J. Hodges, S.J. Press and A. Zellner (Eds.), *Bayesian and Likelihood Methods in Statistics and Econometrics*, North-Holland, Amsterdam, 473-488.

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.

Morimune, K. (1983). "Approximate Distributions of k-Class Estimators When the Degree of Overidentifiability Is Large Compared with the Sample Size," *Econometrica,* 51, 821-841.

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. (1983). "Exact Small Sample Theory in the Simultaneous Equations Model," In: Intriligator,M.D.,Griliches,Z.(Eds.), *Handbook of Econometrics*, North-Holland, Amsterdam, 449-516.

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.

Phillips, P.C.B. (1998a). "New Tools for Understanding Spurious Regressions," *Econometrica*, 66, 1299-1325.

Phillips, P.C.B. (1998b). "New Unit Root Asymptotics in the Presence of Deterministic Trends," Cowles Foundation Discussion Paper No. 1196.

Phillips, P.C.B. and H. R. Moon (1999). "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica,* 67, 1057-1111.

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.

Ullah, A. (1974). "On the Sampling Distributions of Improved Estimators for Coefficients in Linear Regression," *Journal of Econometrics*, 2, 143-150.

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

Table 1: Regression Results Based on Bias Approximations[1]

Dependent Variable $-->$ Regressor	$IV Bias$			$IV Bias/OLS Bias$		
$intercept$	0.003	-0.056	-1.2E-7	0.002	-0.022	0.599
	(66.42)	(-112.2)	(-1.2E-5)	(211.9)	(-246.8)	(640.3)
$\mu'\mu/k_{21}$	–	–	–	–	–	-0.047
	–	–	–	–	–	(-195.1)
$\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho$	–	–	0.599	–	–	–
	–	–	(342.3)	–	–	–
$\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho(\mu'\mu/k_{21})$	–	–	-0.047	–	–	–
	–	–	(-195.1)	–	–	–
$\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho(1+\mu'\mu/k_{21})^{-1}$	1.000	1.006	–	–	–	–
	(76860)	(6885)	–	–	–	–
$\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho(1+\mu'\mu/k_{21})^{-1}/k_{21}$	-2.244	–	–	–	–	–
	(-2347)	–	–	–	–	–
$\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho(1+\mu'\mu/k_{21})^{-2}/k_{21}$	4.561	–	–	–	–	–
	(1300)	–	–	–	–	–
$\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho(1+\mu'\mu/k_{21})^{-3}/k_{21}$	-2.320	–	–	–	–	–
	(-841.3)	–	–	–	–	–
$(1+\mu'\mu/k_{21})^{-1}$	–	–	–	0.997	1.028	–
	–	–	–	(53651)	(6427)	–
$(1+\mu'\mu/k_{21})^{-1}/k_{21}$	–	–	–	-2.237	–	–
	–	–	–	(-2455)	–	–
$(1+\mu'\mu/k_{21})^{-2}/k_{21}$	–	–	–	4.850	–	–
	–	–	–	(1587)	–	–
$(1+\mu'\mu/k_{21})^{-3}/k_{21}$	–	–	–	-2.471	–	–
	–	–	–	(-1104)	–	–
$\bar{R}^2$	1.0000	0.9989	0.7165	1.0000	0.9988	0.4274

[1] Notes: Actual bias and MSE values are generated using the analytical formulae given in Theorems 3.1 and 3.2, as discussed above. Regressions with these values as dependent variable and with various regressors corresponding to the approximations given in Theorem 4.1 and in Section 6 are reported on, with t-statistics in parentheses. Note that in the cannonical model for which our analytical values are generated, least squares bias is equal to $\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho$. The regressor given in the table as $\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho(1+\mu'\mu/k_{21})^{-1}$ is the first term in the approximation. The three regressors subsequent to this one in the table are higher order terms in the expansion. Hence, the regression reported in the first column of entries corresponds to an regression which is based on an approximation which uses only first order terms from the expansion, while the second column of entries reports a regression additionally based on the use of second order terms from the expansion (see above for further details). All regressions reported on in this table use IV Bias and relative bias as dependent variable.

Table 2: Regression Results Based on MSE Approximations[1]

Dependent Variable $-->$ Regressor	$IVMSE$			$IVMSE/OLSMSE$		
intercept	-0.076	0.029	-0.299	-0.008	0.004	0.396
	(-48.71)	(16.33)	(-3.628)	(-54.43)	(19.71)	(345.8)
$\sigma_{uu}\sigma_{vv}^{-1}\rho^2$	–	–	0.396	–	–	
	–	–	(238.9)	–	–	–
$\sigma_{uu}\sigma_{vv}^{-1}\rho^2(\mu'\mu/k_{21})$	–	–	-0.038	–	–	–
	–	–	(-104.7)	–	–	–
$(1-\rho^2)\rho^{-2}/k_{21}$	–	–	–	–		0.298
	–	–	–	–	–	(17.98)
$\mu'\mu/k_{21}$	–	–	–	–	–	-0.038
	–	–	–	–	–	(-126.1)
$\sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)/k_{21}$	–	–	8.059	–	–	–
	–	–	(7.543)	–	–	–
$(1+\mu'\mu/k_{21})^{-2}$	–	–	–	1.002	1.001	–
	–	–	–	(3037)	(1889)	–
$\sigma_{uu}\sigma_{vv}^{-1}\rho^2(1+\mu'\mu/k_{21})^{-2}$	0.996	0.999	–	–	–	–
	(10919)	(10952)	–	–	–	–
$\sigma_{uu}\sigma_{vv}^{-1}(1+\mu'\mu/k_{21})^{-1}/k_{21}$	1.285	–	–	–	–	–
	(139.6)	–	–	–	–	–
$\rho^{-2}(1+\mu'\mu/k_{21})^{-1}/k_{21}$	–	–	–	1.495	–	–
	–	–	–	(438.2)	–	–
$(1+\mu'\mu/k_{21})^{-2}/k_{21}$	–	–	–	-5.684	–	–
	–	–	–	(-70.48)	–	–
$(1+\mu'\mu/k_{21})^{-3}/k_{21}$	–	–	–	8.679	–	–
	–	–	–	(40.98)	–	–
$(1+\mu'\mu/k_{21})^{-4}/k_{21}$	–	–	–	-4.261	–	–
	–	–	–	(-31.45)	–	–
$\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho^2(1+\mu'\mu/k_{21})^{-2}/k_{21}$	-6.370	–	–	–	–	–
	(-100.3)	–	–	–	–	–
$\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho^2(1+\mu'\mu/k_{21})^{-3}/k_{21}$	11.13	–	–	–	–	–
	(87.69)	–	–	–	–	–
$\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho^2(1+\mu'\mu/k_{21})^{-4}/k_{21}$	-5.979	–	–	–	–	–
	(-81.13)	–	–	–	–	–
$\bar{R}^2$	0.9998	0.9996	0.5494	0.9974	0.9859	0.2392

[1] Notes: See notes to Table 1. All regressions reported on in this table use IV Bias and relative bias as dependent variable.

Table 3: Bias of OLS, IV, and Bias Adjusted Estimators[1]

β	$\overline{\pi}$	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{IV}-\beta$	$\widetilde{\beta}_{OLS,1}-\beta$	$\widetilde{\beta}_{OLS,2}-\beta$	$\widetilde{\beta}_{IV}-\beta$	$\widetilde{\widetilde{\beta}}_{IV,1}-\beta$	$\widetilde{\widetilde{\beta}}_{IV,2}-\beta$
-0.5000	0.0500	0.4441	0.2222	-0.0371	0.0569	0.0183	0.0209	0.0591
-0.5000	0.0750	0.3899	0.1304	-0.0978	0.0461	0.0117	0.0144	0.0480
-0.5000	0.1000	0.3331	0.0826	-0.1591	0.0369	0.0073	0.0095	0.0382
-0.5000	0.1250	0.2805	0.0561	-0.2140	0.0300	0.0049	0.0066	0.0308
-0.5000	0.1500	0.2352	0.0404	-0.2607	0.0246	0.0035	0.0048	0.0251
-1.0000	0.0500	0.8899	0.4428	-0.0775	0.1119	0.0341	0.0394	0.1161
-1.0000	0.0750	0.7814	0.2600	-0.1982	0.0910	0.0219	0.0273	0.0947
-1.0000	0.1000	0.6675	0.1647	-0.3203	0.0731	0.0137	0.0180	0.0757
-1.0000	0.1250	0.5622	0.1119	-0.4300	0.0595	0.0092	0.0125	0.0612
-1.0000	0.1500	0.4714	0.0805	-0.5234	0.0488	0.0066	0.0091	0.0499
-1.5000	0.0500	1.3351	0.6635	-0.1186	0.1658	0.0489	0.0567	0.1721
-1.5000	0.0750	1.1724	0.3897	-0.2981	0.1356	0.0319	0.0400	0.1413
-1.5000	0.1000	1.0017	0.2469	-0.4808	0.1093	0.0200	0.0264	0.1132
-1.5000	0.1250	0.8438	0.1678	-0.6452	0.0889	0.0135	0.0184	0.0915
-1.5000	0.1500	0.7075	0.1206	-0.7853	0.0731	0.0096	0.0134	0.0747
-2.0000	0.0500	1.7799	0.8844	-0.1602	0.2189	0.0628	0.0733	0.2273
-2.0000	0.0750	1.5632	0.5195	-0.3978	0.1802	0.0419	0.0527	0.1879
-2.0000	0.1000	1.3357	0.3291	-0.6409	0.1455	0.0263	0.0350	0.1507
-2.0000	0.1250	1.1252	0.2238	-0.8599	0.1185	0.0178	0.0244	0.1219
-2.0000	0.1500	0.9436	0.1609	-1.0466	0.0974	0.0127	0.0178	0.0995
-2.5000	0.0500	2.2246	1.1054	-0.2020	0.2716	0.0763	0.0894	0.2823
-2.5000	0.0750	1.9539	0.6493	-0.4973	0.2249	0.0520	0.0655	0.2343
-2.5000	0.1000	1.6696	0.4114	-0.8006	0.1815	0.0327	0.0436	0.1880
-2.5000	0.1250	1.4066	0.2796	-1.0744	0.1480	0.0221	0.0304	0.1522
-2.5000	0.1500	1.1796	0.2010	-1.3077	0.1216	0.0158	0.0222	0.1243
-3.0000	0.0500	2.6693	1.3264	-0.2442	0.3241	0.0897	0.1054	0.3369
-3.0000	0.0750	2.3446	0.7791	-0.5967	0.2694	0.0619	0.0781	0.2809
-3.0000	0.1000	2.0036	0.4937	-0.9605	0.2178	0.0391	0.0521	0.2256
-3.0000	0.1250	1.6880	0.3356	-1.2886	0.1774	0.0263	0.0363	0.1825
-3.0000	0.1500	1.4156	0.2412	-1.5686	0.1458	0.0188	0.0265	0.1491
-3.5000	0.0500	3.1140	1.5477	-0.2865	0.3765	0.1028	0.1211	0.3914
-3.5000	0.0750	2.7352	0.9092	-0.6961	0.3140	0.0719	0.0908	0.3274
-3.5000	0.1000	2.3375	0.5760	-1.1201	0.2539	0.0456	0.0605	0.2630
-3.5000	0.1250	1.9694	0.3915	-1.5030	0.2070	0.0307	0.0423	0.2129
-3.5000	0.1500	1.6516	0.2814	-1.8295	0.1701	0.0220	0.0309	0.1739
-4.0000	0.0500	3.5586	1.7686	-0.3288	0.4288	0.1160	0.1370	0.4459
-4.0000	0.0750	3.1259	1.0390	-0.7952	0.3585	0.0819	0.1035	0.3739
-4.0000	0.1000	2.6714	0.6582	-1.2799	0.2901	0.0518	0.0692	0.3005
-4.0000	0.1250	2.2508	0.4475	-1.7174	0.2366	0.0351	0.0482	0.2431
-4.0000	0.1500	1.8876	0.3216	-2.0903	0.1944	0.0251	0.0353	0.1988
-4.5000	0.0500	4.0033	1.9900	-0.3711	0.4812	0.1289	0.1525	0.5000
-4.5000	0.0750	3.5166	1.1689	-0.8946	0.4034	0.0918	0.1160	0.4202
-4.5000	0.1000	3.0053	0.7405	-1.4395	0.3261	0.0583	0.0775	0.3378
-4.5000	0.1250	2.5322	0.5034	-1.9316	0.2659	0.0392	0.0542	0.2735
-4.5000	0.1500	2.1237	0.3619	-2.3514	0.2187	0.0282	0.0395	0.2236
-5.0000	0.0500	4.4479	2.2111	-0.4132	0.5333	0.1419	0.1681	0.5547
-5.0000	0.0750	3.9073	1.2988	-0.9940	0.4478	0.1020	0.1289	0.4669
-5.0000	0.1000	3.3393	0.8228	-1.5992	0.3623	0.0644	0.0862	0.3753
-5.0000	0.1250	2.8136	0.5593	-2.1459	0.2954	0.0436	0.0602	0.3039
-5.0000	0.1500	2.3597	0.4021	-2.6122	0.2430	0.0311	0.0439	0.2485

[1] Notes: The first and second columns report values of β and $\overline{\pi}$ used to parameterize the DGP. The 3rd column reports the bias of the OLS estimator, the 4th column reports the bias of the IV estimator, and the last 6 columns report the bias of various bias adjusted estimators of β. All numerical bias entries are averages across 5000 Monte Carlo trials.

Table 4: MSE of OLS, IV, and Bias Adjusted Estimators[1]

β	$\overline{\pi}$	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{IV}-\beta$	$\widetilde{\beta}_{OLS,1}-\beta$	$\widetilde{\beta}_{OLS,2}-\beta$	$\widetilde{\beta}_{IV}-\beta$	$\widetilde{\widetilde{\beta}}_{IV,1}-\beta$	$\widetilde{\widetilde{\beta}}_{IV,2}-\beta$
-0.5000	0.0500	0.1990	0.0591	0.0479	0.0365	0.0403	0.0400	0.0365
-0.5000	0.0750	0.1536	0.0228	0.0265	0.0124	0.0123	0.0122	0.0124
-0.5000	0.1000	0.1123	0.0105	0.0353	0.0064	0.0059	0.0059	0.0065
-0.5000	0.1250	0.0799	0.0057	0.0530	0.0039	0.0035	0.0035	0.0040
-0.5000	0.1500	0.0563	0.0035	0.0737	0.0027	0.0023	0.0023	0.0027
-1.0000	0.0500	0.7939	0.2084	0.0820	0.0713	0.0721	0.0719	0.0718
-1.0000	0.0750	0.6125	0.0755	0.0639	0.0255	0.0211	0.0211	0.0260
-1.0000	0.1000	0.4474	0.0324	0.1163	0.0136	0.0100	0.0100	0.0139
-1.0000	0.1250	0.3177	0.0163	0.1947	0.0084	0.0058	0.0058	0.0086
-1.0000	0.1500	0.2237	0.0093	0.2819	0.0057	0.0038	0.0038	0.0057
-1.5000	0.0500	1.7847	0.4565	0.1435	0.1322	0.1292	0.1292	0.1337
-1.5000	0.0750	1.3769	0.1630	0.1267	0.0475	0.0362	0.0362	0.0486
-1.5000	0.1000	1.0057	0.0689	0.2511	0.0255	0.0168	0.0168	0.0262
-1.5000	0.1250	0.7142	0.0339	0.4302	0.0159	0.0096	0.0097	0.0163
-1.5000	0.1500	0.5027	0.0189	0.6278	0.0107	0.0063	0.0063	0.0109
-2.0000	0.0500	3.1707	0.8038	0.2340	0.2207	0.2137	0.2137	0.2232
-2.0000	0.0750	2.4466	0.2855	0.2146	0.0783	0.0575	0.0577	0.0805
-2.0000	0.1000	1.7872	0.1199	0.4395	0.0424	0.0263	0.0265	0.0436
-2.0000	0.1250	1.2693	0.0586	0.7590	0.0263	0.0151	0.0151	0.0270
-2.0000	0.1500	0.8933	0.0323	1.1109	0.0176	0.0097	0.0098	0.0180
-2.5000	0.0500	4.9520	1.2499	0.3546	0.3372	0.3258	0.3263	0.3417
-2.5000	0.0750	3.8215	0.4433	0.3279	0.1181	0.0848	0.0852	0.1215
-2.5000	0.1000	2.7919	0.1854	0.6812	0.0639	0.0385	0.0388	0.0659
-2.5000	0.1250	1.9829	0.0904	1.1812	0.0399	0.0220	0.0222	0.0408
-2.5000	0.1500	1.3956	0.0497	1.7314	0.0267	0.0142	0.0144	0.0272
-3.0000	0.0500	7.1288	1.7957	0.5061	0.4833	0.4669	0.4676	0.4890
-3.0000	0.0750	5.5018	0.6360	0.4660	0.1667	0.1188	0.1194	0.1717
-3.0000	0.1000	4.0198	0.2657	0.9764	0.0900	0.0535	0.0539	0.0929
-3.0000	0.1250	2.8551	0.1291	1.6968	0.0562	0.0305	0.0305	0.0577
-3.0000	0.1500	2.0095	0.0709	2.4890	0.0376	0.0197	0.0199	0.0384
-3.5000	0.0500	9.7010	2.4408	0.6888	0.6577	0.6374	0.6383	0.6665
-3.5000	0.0750	7.4875	0.8641	0.6297	0.2239	0.1587	0.1595	0.2306
-3.5000	0.1000	5.4709	0.3605	1.3251	0.1215	0.0711	0.0717	0.1253
-3.5000	0.1250	3.8860	0.1749	2.3060	0.0754	0.0404	0.0408	0.0777
-3.5000	0.1500	2.7351	0.0960	3.3839	0.0506	0.0263	0.0263	0.0517
-4.0000	0.0500	12.668	3.1862	0.9033	0.8627	0.8374	0.8387	0.8729
-4.0000	0.0750	9.7786	1.1265	0.8175	0.2904	0.2044	0.2054	0.2992
-4.0000	0.1000	7.1452	0.4694	1.7270	0.1572	0.0922	0.0922	0.1622
-4.0000	0.1250	5.0754	0.2279	3.0082	0.0980	0.0523	0.0523	0.1005
-4.0000	0.1500	3.5724	0.1247	4.4162	0.0654	0.0336	0.0339	0.0668
-4.5000	0.0500	16.032	4.0289	1.1495	1.0966	1.0677	1.0693	1.1094
-4.5000	0.0750	12.375	1.4244	1.0321	0.3663	0.2574	0.2586	0.3774
-4.5000	0.1000	9.0428	0.5941	2.1829	0.1980	0.1148	0.1157	0.2044
-4.5000	0.1250	6.4235	0.2878	3.8040	0.1233	0.0655	0.0655	0.1265
-4.5000	0.1500	4.5213	0.1578	5.5861	0.0823	0.0420	0.0425	0.0841
-5.0000	0.0500	19.790	4.9735	1.4269	1.3596	1.3280	1.3299	1.3774
-5.0000	0.0750	15.277	1.7584	1.2711	0.4507	0.3162	0.3178	0.4644
-5.0000	0.1000	11.163	0.7323	2.6916	0.2434	0.1407	0.1418	0.2512
-5.0000	0.1250	7.9303	0.3553	4.6939	0.1515	0.0801	0.0801	0.1562
-5.0000	0.1500	5.5819	0.1943	6.8931	0.1010	0.0514	0.0519	0.1038

[1] Notes: See notes to Table 3. The 3rd column reports the MSE of the OLS estimator, the 4th column reports the MSE of the IV estimator, and the last 6 columns report the MSE of various bias adjusted estimators of β, relative to the OLS MSE.

ftbpF4.0274in3.4982in0ptgr15a.wmf

ftbpF4.6276in3.4333in0ptgr15b.wmf

ftbpF6.0269in3.3736in0ptbgmg5a.wmf

ftbpF6.0269in3.3537in0ptbgmg5b.wmf