On The Bias and MSE of The IV Estimator Under Weak Identification[*]

John Chao — University of Maryland
Norman R. Swanson — Texas A&M University

January 2000

Abstract

In this paper we provide further results on the properties of the IV estimator in the presence of weak instruments.We begin by formalizing the notion of weak identification within the local-to-zero asymptotic framework of Staiger and Stock (1997), and deriving explicit analytical formulae for the asymptotic bias and mean square error (MSE) of the IV estimator. These results generalize earlier findings by Staiger and Stock (1997), who give an approximate measure for the asymptotic bias of the two-stage least squares (2SLS) estimator relative to that of the OLS estimator. We also show that in the special case where all available instruments are used and where the underlying simultaneous equations model has an orthonormal canonical structure, the bias and MSE formulae which we obtain are identical to the exact bias and MSE of the 2SLS estimator obtained by Richardson and Wu (1971) under Gaussian error assumptions. This result gives a partial confirmation to the Staiger-Stock assertion, based on intuitive arguments, that the limiting distribution of the 2SLS estimator derived under the more general assumptions of the Staiger-Stock local-to-zero asymptotic framework coincides with the exact distribution of the same estimator derived under the more restrictive assumptions of a fixed instrument/Gaussian model. Because our analytical formulae for bias and MSE are complex functionals of confluent hypergeometric functions, we also derive approximations for these formulae which are based on an expansion that allows the number of instruments to grow to infinity while keeping the population analogue of the first stage F-statistic fixed. In addition, we provide a series of regression results that show this expansion to give excellent approximations for the bias and MSE functions in general. These approximations allow us to make several 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 asymptotic bias 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. In addition, 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 relative to the bias component in a scenario where the number of instruments used is taken to be large while the population analogue of the first stage F-statistic is kept constant. One feature of our approach which ties our findings to the earlier IV literature is that our results apply not only to the weak instrument case asymptotically, but also to the finite sample case with fixed (possibly good) instruments and Gaussian errors, since our formulae correspond to the exact bias and MSE functionals when a fixed instrument/Gaussian model is assumed.

JEL classification: C12, C22.
Keywords: local to zero asymptotics, confluent hypergeometric function.

---

[*]John Chao: Department of Economics, University of Maryland, College Park, MD, USA 20742, chao@econ.umd.edu. Norman R. Swanson: Department of Economics, Texas A&M University, College Station, TX, USA 77843-4228, nswanson@econ.tamu.edu. The authors wish to thank Carmela Quintos for helpful comments. In addition, Swanson thanks the Private Enterprise Research Center at Texas A&M for financial support.

1 Introduction

There has been much recent interest in instrumental variables regression with instruments that are only weakly correlated with the endogenous explanatory variables. Important theoretical contributions to this expanding literature include such papers as Nelson and Startz (1990a), Dufour (1996), Staiger and Stock (1997), and Zivot and Wang (1998). Much of this literature focuses on the impact that using weak instruments has on interval estimation and hypothesis testing. In contrast, fewer results have been obtained characterizing the properties of point estimators under weak identification. This is despite the fact that empirical researchers who first noted the problem of weak instruments were clearly very interested in the consequences for point estimation, as can be seen from the papers by Nelson and Startz (1990b), Bound, Jaeger, and Baker (1995), and Angrist and Krueger (1995). For example, Bound, Jaeger, and Baker (1995) point out that while it has been widely recognized that the use of weak instruments leads instrumetal variable (IV) estimators with higher standard errors. They also point out that weak instruments can also lead to larger bias relative to least squares - an observation which in large part has accounted for the renewed interest in the problem.

The purpose of this paper is to provide further results on the properties of the IV estimator in the presence of weak instruments. We begin our analysis with a standard single-equation simultaneous equation setup, where the set of available instruments may be arbitrarily large. In order to keep notation manageable, we allow the single structural equation of interest to have an arbitrary number of exogenous explanatory variables but only one endogenous explanatory variable[1]. In addition, we adopt the local-to-zero asymptotic framework of Staiger and Stock (1997) to formalize the notion of weak identification. One interesting feature of the Staiger and Stock setup is that they show (via a series of simulations) that their framework yields a very good approximation for the finite sample distribution of the IV estimator when the quality of the available instruments is poor. In contrast to Staiger and Stock (1997), who give an approximate measure for the asymptotic bias of the two-stage least squares (2SLS) estimator relative to that of the OLS estimator, we derive explicit analytical formulae for the asymptotic bias and mean square error (MSE) of a more general IV estimator. These formulae can be used to characterize 2SLS estimator bias and MSE

[1] Although we only study the case with one endogenous explanatory variable, many of the qualitative conclusions reached in this paper can be generalized in a straightforward manner to more general settings.

when all available instruments are used. We also show that in the special case where all available instruments are used and where the underlying simultaneous equations model has an orthonormal canonical structure (to be defined below), the bias and MSE formulae which we obtain are identical to the exact bias and MSE of the 2SLS estimator obtained by Richardson and Wu (1971) under Gaussian error assumptions. This result gives a partial confirmation to the Staiger-Stock assertion, based on intuitive arguments, that the limiting distribution of the 2SLS estimator derived under the more general assumptions of the Staiger-Stock local-to-zero asymptotic framework coincides with the exact distribution of the same estimator derived under the more restrictive assumptions of a fixed instrument/Gaussian model.

Because the analytical formulae for bias and MSE are complex functionals of confluent hypergeometric functions, we also derive approximations for these formulae which are based on an expansion that allows the number of instruments to grow to infinity while keeping the population analogue of the first stage F-statistic fixed. To the best of our knowledge, the expansion which we use is new to the IV literature. In addition, we provide a series of regression results that show this expansion to give excellent approximations for the bias and MSE functions in general. Moreover, the approximations allow us to make several interesting additional observations. First, when the approximation method is applied to the bias, the lead term of the expansion, when appropriately standardized by the asymptotic bias 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. Second, 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 relative to the bias component in a scenario where the number of instruments used is taken to be large while the population analogue of the first stage F-statistic is kept constant. Third, our approximations show that even when all available instruments are weak in the sense of Staiger and Stock (1997), increasing the number of instruments used will still reduce the MSE by reducing the variance of estimation, although the magnitude of this reduction becomes small as the number of instruments used becomes large. We note, finally, that our results not only apply to the weak instrument case asymptotically but they also apply to the finite sample case with fixed, and possibly good, instruments and Gaussian errors, since our formulae correspond to the exact bias and MSE functionals when a fixed instrument/Gaussian model is assumed.

This rest of the paper is organized as follows. Section 2 contains preliminaries, including the model, assmuptions, and notation to be used. Section 3 presents formulae for the asymptotic bias and MSE of the IV estimator under weak identification, and discusses properties and implications of the formulae. Section 4 develops bias and MSE approximations. Section 5 summarizes various numerical calculations and regression findings based on our exact and approximate bias and MSE formulae. Conclusions and summarizing remarks are given in Section 6. All proofs and technical details are contained in the Appendix.

2 Setup

Consider the simultaneous equations model (SEM):

$$y_1 = y_2\beta + X\gamma + u; \tag{1}$$

$$y_2 = Z\Pi + X\phi + v; \tag{2}$$

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 of the k_1 exogenous variables included in the structural equation (1), Z is an $T \times k_2$ matrix of observations of the k_2 exogenous variables excluded from the structural equation (1), and u and v are each $T \times 1$ vector of random disturbances. Let u_t and v_t devote the t-th component of the random vectors u and v, respectively. We assume $w_t = (u_t; v_t)'$ to be white noise so that $E(w_t) = 0$; $E(w_t w_s') = 0$ for all $t \neq s$, and $E(w_t w_t') = \Sigma$, where we can partition the 2×2 matrix Σ conformably with $(u_t; v_t)'$ as

$$\Sigma = \begin{pmatrix} \sigma_{uu} & \sigma_{vu} \\ \sigma_{vu} & \sigma_{vv} \end{pmatrix} : \tag{3}$$

Following Staiger and Stock (1997), we formalize the case of weak instruments by modeling Π to be a parameter sequence that is local to zero, i.e.

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 processes of the exogenous variables $\overline{Z} = (X; Z)$ and of the disturbances $(u; v)$ in equation (1) and (2) above are such that the following moment covergence results hold.

Assumption 2: The following limits hold jointly: (i) $(u'u/T; v'u/T; v'v/T) \xrightarrow{p} (\sigma_{uu}; \sigma_{vu}; \sigma_{vv})$; (ii) $\overline{Z}'\overline{Z}/T \xrightarrow{p} Q$; and (iii) $(T^{-1/2}X'u; T^{-1/2}Z'u; T^{-1/2}X'v; T^{-1/2}Z_v') \Rightarrow (\psi_{Xu}; \psi_{Zu}; \psi_{Xv}; \psi_{Zv})$;

where $Q = E(\overline{Z}_t\overline{Z}_t')$ with $\overline{Z}_t$ the t-th row of the matrix $\overline{Z}$; and where $\psi \equiv (\psi_{Xu}', \psi_{Zu}', \psi_{Xv}', \psi_{Zv}')'$ is distributed $N(\underline{0}, (\Sigma \otimes Q))$.

In this paper, 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

$$\hat{\beta}_{IV} = (y_2'(P_H - P_X)y_2)^{-1}(y_2'(P_H - P_X)y_1), \tag{4}$$

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. To introduce som notations, we find it 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 in 4 above is equivalent to the 2SLS estimator. Moreover, to further fix notations, we 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 $\Pi_T = [\Pi_{1,T}', \Pi_{2,T}']' = [C_1'/\sqrt{T}, C_2'/\sqrt{T}]'$, $T^{-\frac{1}{2}}Z'u = [T^{-\frac{1}{2}}(Z_1'u)', T^{-\frac{1}{2}}(Z_2'u)']'$, $T^{-\frac{1}{2}}Z'v = [T^{-\frac{1}{2}}(Z_1'v)', T^{-\frac{1}{2}}(Z_2'v)']'$, and $\psi_{Zu} = [\psi_{Z_1u}', \psi_{Z_2u}']'$, and $\psi_{Zv} = [\psi_{Z_1v}', \psi_{Z_2v}']'$, where from assumption 2 part (iii) above, we have, of course, that $(T^{-\frac{1}{2}}(Z_1'u)', T^{-\frac{1}{2}}(Z_2'u)', T^{-\frac{1}{2}}(Z_1'v)', T^{-\frac{1}{2}}(Z_2'v)') \Rightarrow (\psi_{Z_1u}', \psi_{Z_2u}', \psi_{Z_1v}', \psi_{Z_2v}')'$. In addition, we partition Q conformably with $\overline{Z} = (X, Z_1, Z_2)$ as

$$Q = \begin{pmatrix} Q_{XX} & Q_{XZ_1} & Q_{XZ_2} \\ Q_{Z_1X} & Q_{Z_1Z_1} & Q_{Z_1Z_2} \\ Q_{Z_2X} & Q_{Z_2Z_1} & Q_{Z_2Z_2} \end{pmatrix}. \tag{5}$$

Finally, we define

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12}' & \Omega_{22} \end{pmatrix} \tag{6}$$

$$= \begin{pmatrix} Q_{Z_1Z_1} - Q_{Z_1X}Q_{XX}^{-1}Q_{XZ_1} & Q_{Z_1Z_2} - Q_{Z_1X}Q_{XX}^{-1}Q_{XZ_2} \\ Q_{Z_2Z_1} - Q_{Z_2X}Q_{XX}^{-1}Q_{XZ_1} & Q_{Z_2Z_2} - Q_{Z_2X}Q_{XX}^{-1}Q_{XZ_2} \end{pmatrix}. \tag{7}$$

and

$$\Omega_{*1} = (\Omega_{11}, \Omega_{12}). \tag{8}$$

To ensure that the asymptotic bias and MSE of the IV estimator are well-behaved, we make the further assumption

Assumption 3: $\sup_T E[|W_T|^{2+\varepsilon}] < \infty$; for some $\varepsilon > 0$; where $W_T = \hat{\beta}_{IV,T} - \beta_0$ and where $\hat{\beta}_{IV,T}$ 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 $(\hat{\beta}_{IV,T} - \beta_0)^2$ (see Billingsley (1968), pp.32). Under assumption 3, $\lim_{T\to\infty} E(\hat{\beta}_{IV,T} - \beta_0) = E(W)$ and $\lim_{T\to\infty} E(\hat{\beta}_{IV,T} - \beta_0)^2 = E(W^2)$, where W is the limiting random variable of the sequence $\{W_T\}$ whose explicit form is given in Lemma A1 in the Appendix. Hence, under assumption 3, the asymptotic bias and MSE correspond with bias and MSE as implied by the limiting distribution of $\hat{\beta}_{IV,T}$. 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 IV estimator of β under Gaussian assumptions has finite sample moments which exist up to and including the degree of apparent overidentification, as given by the order condition (see, for example, Sawa (1969)). Throughout this paper, we shall assume $k_2 \geq 4$ so as to ensure that our results will apply in the Gaussian case.

3 Analytical Results

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

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

(a)

$$b_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho e^{-\frac{\mu'\mu}{2}} {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right); \qquad (9)$$

where $b_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) = \lim_{T\to\infty} E(\hat{\beta}_{IV,T} - \beta_0)$ 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}; $\rho = \sigma_{uv}\sigma_{uu}^{-\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}$, $\Gamma(\cdot)$ denotes the gamma function, and ${}_1F_1(\cdot;\cdot;\cdot)$ denotes the confluent hypergeometric function.

(b) For k_{21} fixed, as $\mu'\mu \to \infty$;

$$b_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) \to 0. \qquad (10)$$

(c) For $\mu'\mu$ fixed, as $k_{21} \to \infty$;

$$b_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) \to \sigma_{vu}/\sigma_{vv} \tag{11}$$

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

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

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

(a)

$$\begin{aligned} m_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) \;=\; & \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}} \times \\ & \left[\frac{1}{\rho^2}\left(\frac{1}{k_{21}-2}\right) {}_1F_1\left(\frac{k_{21}}{2}-1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right) \right. \\ & \left. + \left(\frac{k_{21}-3}{k_{21}-2}\right) {}_1F_1\left(\frac{k_{21}}{2}-2; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right)\right]; \end{aligned} \tag{12}$$

where $m_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) = \lim_{T\to\infty} E\left(\hat{\beta}_{IV,T} - \beta_0\right)^2$ 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 above.

(b) For k_{21} fixed, as $\mu'\mu \to \infty$;

$$m_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) \to 0. \tag{13}$$

(c) For $\mu'\mu$ fixed, as $k_{21} \to \infty$;

$$m_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) \to \sigma_{uu}\sigma_{vv}^{-1}\rho^2. \tag{14}$$

(d) The asymptotic mean squared error function $m_{\hat{\beta}_{IV}}(\mu'\mu; k_{21})$ is a monotonically decreasing function of $\mu'\mu$ for k_{21} fixed.

Remark 3.3:

(i) Note that the asymptotic bias and MSE formulae, given by expressions (9) and (12), are functionals of 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 (see, for example, Slater (1960), pp.2) so that

$$ {}_1F_1(a;\ b;\ x) = \sum_{j=0}^{\infty} \frac{(a)_j}{(b)_j}\frac{x^j}{j!} \tag{15} $$

where the notation $(a)_j$ denotes the Pochhammer's symbol, i.e.,

$$ (a)_j = \begin{cases} \Gamma(a+j)/\Gamma(a) & \text{for integer } j \geq 1 \\ 1 & \text{for } j = 0 \end{cases}. \tag{16} $$

It follows that the bias expression (9) and the MSE expression (12) can both be written in infinite series form, i.e.

$$ b_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho e^{-\frac{\mu'\mu}{2}} \left[\sum_{j=0}^{\infty} \frac{(\frac{k_{21}}{2}-1)_j}{(\frac{k_{21}}{2})_j} \frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!} \right]; \tag{17} $$

$$ \begin{aligned} m_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}} \times \\ &\left[\frac{1}{\rho^2}\left(\frac{1}{k_{21}-2}\right) \sum_{j=0}^{\infty} \frac{(\frac{k_{21}}{2}-1)_j}{(\frac{k_{21}}{2})_j} \frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!} \right. \\ &\left. + \left(\frac{k_{21}-3}{k_{21}-2}\right) \sum_{j=0}^{\infty} \frac{(\frac{k_{21}}{2}-2)_j}{(\frac{k_{21}}{2})_j} \frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!} \right]. \end{aligned} \tag{18} $$

The main merits of the infinite series representations given in expressions (17) and (18) lie in the fact that they provide explicit formulae for the asymptotic bias and MSE of the IV estimator under weak indentification. We have, in fact, made extensive uses of these representations in deriving some of the properties of the asymptotic bias and MSE, as reported in Theorems 3.1 and 3.2 above. (For details, see the proof of these theorems as presented in the Appendix of this paper.) These infinite series representations are, however, quite complicated and difficult to analyze. Hence, in order to give more tractable representations, we shall report in the next section some approximations which yield much simpler expressions for the asymptotic bias and MSE.

(ii) From part (b) of Theorem 3.2, we see that the MSE function for $\hat{\beta}_{IV;T}$ goes to zero as $\mu'\mu \to \infty$. Note that the case where $\mu'\mu \to \infty$ corresponds roughly with the case where the available instruments are not weak but fully relevant. In this case, then, Theorem 3.2 part (b) shows $\hat{\beta}_{IV;T}$

to converge in a mean squared sense to the true value β_0. It follows, of course, that in this case $\hat{\beta}_{IV,T}$ is a (weakly) consistent estimator of β, as one would expect from a conventional asymptotic analysis with a fully identified model. Hence, our result can be interpreted as having the textbook case of good instruments as a limiting special case.

(iii) Observe that under a condition similar to Assumption 3 above, it can be shown that

$$\lim_{T \to \infty} E\left(\hat{\beta}_{OLS,T} - \beta_0\right) = \sigma_{uv}/\sigma_{vv}; \tag{19}$$

$$\lim_{T \to \infty} E\left(\hat{\beta}_{OLS,T} - \beta_0\right)^2 = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 = \sigma_{uv}^2/\sigma_{vv}^2; \tag{20}$$

where $\hat{\beta}_{OLS,T}$ is the OLS estimator. To see this, let $W_T^* = \hat{\beta}_{OLS,T} - \beta_0$ and assume that

$$\sup_{T \geq 5} E[|W_T^*|^{2+\epsilon}] < \infty. \tag{21}$$

It is well-known that under Gaussian error assumptions, condition (21) is satisfied since in this case the finite sample distribution of the OLS estimator 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 expressions (19) and (20), we note that Staiger and Stock (1997) has shown that under Assumption 2 above,

$$\hat{\beta}_{OLS,T} - \beta_0 \xrightarrow{p} \sigma_{uv}/\sigma_{vv} \text{ as } T \to \infty. \tag{22}$$

Thus, it follows directly from Theorem 5.4 of Billingsley (1968) that

$$\lim_{T \to \infty} E\left(\hat{\beta}_{OLS,T} - \beta_0\right) = \sigma_{uv}/\sigma_{vv}$$

and

$$\begin{aligned} \lim_{T \to \infty} E\left(\hat{\beta}_{OLS,T} - \beta_0\right)^2 &= \lim_{T \to \infty} E(W_T^{*2}) \\ &= E(\sigma_{uv}/\sigma_{vv})^2 \\ &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2. \end{aligned}$$

Moreover, comparing expressions (19) and (20) 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 asymptotic bias and the asymptotic MSE (AMSE henceforth) of the IV estimator converge to those of the OLS estimator as $k_{21} \to \infty$.

(iv) Write the asymptotic bias function of $\hat{\beta}_{IV}$ as

$$b_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) = \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \rho f(\mu'\mu; k_{21}): \tag{23}$$

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 $\sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \rho = \sigma_{vu}/\sigma_{vv}$ is simply the asymptotic bias of the OLS estimator, we see that the bias of the IV estimator as given by (9) has the same sign as the OLS bias. Moreover, note that

$$\begin{aligned} |b_{\hat{\beta}_{IV}}(\mu'\mu; k_{21})| &= |\sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \rho| f(\mu'\mu; k_{21}) \\ &< |\sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \rho|: \end{aligned}$$

Hence, even with weak instruments, the asymptotic bias of the IV estimator in absolute magnitude is less than that of the OLS estimator for $\mu'\mu \neq 0$ and finite k_{21}, and the former only tends to the OLS bias as $k_{21} \to \infty$. In addition, only when $\mu'\mu = 0$ is the asymptotic bias of the two estimators exactly equal for finite k_{21}. Our result, thus, formalizes the intuitive discussion in Bound, Jaeger, and Baker (1995) and Angrist and Krueger (1995) that, with weak instruments, the IV estimator is biased in the direction of the OLS estimator, with the magnitude of the bias approaching that of the OLS estimator as the R^2 between the instruments and the endogenous explanatory variable approaches zero, i.e., as $\mu'\mu \to 0$.

(v) Note that the AMSE function of $\hat{\beta}_{IV}$ is a much more complicated function of $\mu'\mu$ and k_{21} than the bias function. While the asymptotic bias of $\hat{\beta}_{IV}$ is less than that of $\hat{\beta}_{OLS}$ for all positive real number $\mu'\mu$ and for all values of k_{21} for which the bias exists and while the former only tends to the latter as $k_{21} \to \infty$ for a given $\mu'\mu$, the AMSE of $\hat{\beta}_{IV}$ 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 $\hat{\beta}_{OLS}$. To see this, consider the example where $\mu'\mu = 0$; and note that, in this case, expression (12) becomes .

$$m_{\hat{b}_{IV}}(\mu'\mu = 0; k_{21}) = \sigma_{uu}\sigma_{VV}^{-1}\rho^2 \left[1 + \left(\frac{1-\rho^2}{\rho^2}\right)\left(\frac{1}{k_{21}-2}\right)\right]; \tag{24}$$

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 any given k_{21}, so that as $\mu'\mu$ grows the AMSE of the IV estimator will eventually become smaller than that of the OLS estimator.

(vi) It would be of interest to compare our results with those obtained in the extensive literature on the finite sample properties of IV estimators. See Richardson (1968), Sawa (1969), Richardson and Wu (1971), Holly and Phillips (1979), and Phillips (1980, 1983, 1989) as some of the papers which exemplify this literature. To proceed with such a comparison, we note first that the SEM given by expressions (1) and (2) can alternatively be written in the reduced form

$$y_1 = Z\Gamma_1 + X\Gamma_2 + \eta; \tag{25}$$

$$y_2 = Z\Pi + X\Phi + v; \tag{26}$$

where $\Gamma_1 = \Pi\beta$; $\Gamma_2 = \Phi\beta + \gamma$, and $\eta = u + v\beta$. 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

$$\begin{pmatrix} \eta_t \\ v_t \end{pmatrix} \sim i.i.d.N(0; G); \tag{27}$$

where η_t and v_t denote the t-th coordinate of the $T \times 1$ random vectors η and v, respectively, and where

$$G = \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix}: \tag{28}$$

Now, suppose we consider the case where all available instruments are used, i.e., the case where the IV estimator given by expression (4) is simply the 2SLS estimator. Then, it follows that

$$\begin{aligned} \mu'\mu &= \sigma_{VV}^{-1}C' \cdot C \\ &= \sigma_{VV}^{-1}C'(Q_{ZZ} - Q_{ZX}Q_{XX}^{-1}Q_{XZ})C: \end{aligned} \tag{29}$$

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 by expression (3) earlier, can be written as

$$\sigma_{uu} = g_{11} - 2g_{12}\beta + g_{22}\beta^2; \tag{30}$$

$$\sigma_{uv} = g_{12} - g_{22}\beta; \tag{31}$$

$$\sigma_{vv} = g_{22}. \tag{32}$$

Substituting expressions (30), (31), and (32) into the bias formula (9) and the MSE formula (12), we see, after some simple manipulations, that these formulae can be rewritten as

$$b_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) = -\frac{g_{22}\beta - g_{12}}{g_{22}} e^{-\frac{\mu'\mu}{2}} {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right); \tag{33}$$

$$\begin{aligned} m_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) = \frac{g_{11}g_{22} - g_{12}^2}{g_{22}} & \left(\frac{1}{k_{21} - 2}\right)(1 + \tau^2) e^{-\frac{\mu'\mu}{2}} {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right) \\ & + \left(\frac{k_{21} - 3}{k_{21} - 2}\right)\tau^2 e^{-\frac{\mu'\mu}{2}} {}_1F_1\left(\frac{k_{21}}{2} - 2; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right); \end{aligned} \tag{34}$$

where

$$\tau = \frac{g_{22}\beta - g_{12}}{(g_{11}g_{22} - g_{12}^2)^{\frac{1}{2}}}. \tag{35}$$

Comparing expressions (33) and (34) with equations (3.1) and (4.1) of Richardson and Wu (1971), we see that in this case the formulae for the bias and MSE, which hold asymptotically when the instruments are weak in the sense of Staiger and Stock (1997), 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 (33) and (34) above whereas the sample analogue of the concentration parameter appears in the exact formulae reported in Richardson and Wu (1971). Hence, our bias and MSE results partially confirm the Staiger-Stock conjecture that the limiting distribution of the 2SLS estimator obtain under the local-to-zero asymptotic framework is the same as the exact distribution of this estimator under the assumption of a fixed instrument/Gaussian model.

4 Approximation Results

The bias and MSE functions given in Theorems 3.1 and 3.2, respectively, of the last section are functionals of confluent hypergeometric functions and, thus, have infinite series representations which are very complicated as discussed in Remark 3.3(i) of the last section. In this section, we give approximations for the bias and MSE based on an expansion which holds as we allow k_{21}, the number of instruments used, to grow to infinity while keeping $\frac{\mu'\mu}{k_{21}}$, the population analogue of the first stage F-statistic, fixed. These approximations yield simpler expressions for the bias and MSE than the infinite series representations given by (17) and (18) above. We shall give a motivation for the type of expansions used here in Remark 4.2(i) below; in addition, response surface regression results, which we will present in the next section, shows our approximations to work well even when k_{21} is relatively small. First, however, we summarize our approximation results in the following theorem:

Theorem 4.1 (Approximations): Let $\tau^2 = \frac{\mu'\mu}{k_{21}}$ and reparameterize the bias function (9) and the MSE function (12) in terms of τ^2 and k_{21} so that

$$b_{\hat{\beta}_{IV}}(\tau^2; k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho e^{-\frac{\tau^2 k_{21}}{2}} {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\tau^2 k_{21}}{2}\right); \quad (36)$$

$$\begin{aligned} m_{\hat{\beta}_{IV}}(\tau^2; k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\tau^2 k_{21}}{2}} \times \\ &\left[\frac{1}{\rho^2}\left(\frac{1}{k_{21} - 2}\right) {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\tau^2 k_{21}}{2}\right) \right. \\ &\left. + \left(\frac{k_{21} - 3}{k_{21} - 2}\right) {}_1F_1\left(\frac{k_{21}}{2} - 2; \frac{k_{21}}{2}; \frac{\tau^2 k_{21}}{2}\right)\right]. \end{aligned} \quad (37)$$

Then, as $k_{21} \to \infty$ with τ^2 fixed, the following results hold

(a)

$$\begin{aligned} b_{\hat{\beta}_{IV}}(\tau^2; k_{21}) &= \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho \left[\left(\frac{1}{1+\tau^2}\right) \right. \\ &\left. - \frac{2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2\right] + O(k_{21}^{-2}); \end{aligned} \quad (38)$$

(b)

$$m_{\hat{\beta}_{IV}}(\tau^2; k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 \left(\frac{1}{1+\tau^2}\right)^2$$

$$
\begin{aligned}
&+\sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\\
&+\sigma_{uu}\sigma_{vv}^{-1}\rho^2\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left[1+3\left(\frac{1}{1+\tau^2}\right)\right.\\
&\left.+2\left(\frac{1}{1+\tau^2}\right)^2-6\left(\frac{1}{1+\tau^2}\right)^3\right]+O(k_{21}^{-2}). \qquad (39)
\end{aligned}
$$

Remark 4.2:

(i) 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 to approach infinity to obtain the asymptotic bias and MSE under weak identification and then let k_{21} approach infinity with $\tau^2 = \frac{\mu'\mu}{k_{21}}$ held fixed to obtain the approximate formulae (38) and (39). We conjecture that our sequential limit procedure corresponds to a case where a joint limit is taken with respect to T and k_{21} but where k_{21} grows at a slower rate relative to T so that $k_{21}/T \to 0$ as $T, k_{21} \to \infty$. In this sense, our approach differs from that of Bekker (1994) and Hahn (1997), who (in our notation) look essentially at the case where k_{21} and T grow at the same rate, i.e., the case where $k_{21} \to \infty$, $T \to \infty$; and $k_{21}/T \to \alpha$, with α being a positive constant. It would certainly be of interest to consider an asymptotic framework which combines the approach of Bekker (1995) with the local-to-zero approach of Staiger and Stock (1997), but we feel that this work is best left to future research.

It should be emphasized that our sequential limit approach is designed to mimic a typical empirical situation where the sample size is very large, where the number of instruments is large but of a smaller order of magnitude relative to the sample size, and where the value of the first-stage F-statistic is relatively small. A leading example of such an empirical situation is provided by the well-known study reported in Bound, Jaeger, and Baker (1995), which replicates an important study by Angrist and Krueger (1991) on estimating the returns to schooling using quarter of birth as instruments. In that study, the sample size is 329,509, the number of instruments is 180, and the value of the first stage F-statistic is 2.428. We expect our approach to give good approximations for the bias and MSE of the IV estimator in such situations.

(ii) Set

$$
\mathbf{b}_{IV}(\tau^2; k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left[\left(\frac{1}{1+\tau^2}\right)\right.
$$

$$
-\frac{2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2\Bigg]; \tag{40}
$$

so that the notation $b_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ denotes our approximation for the asymptotic bias of the IV estimator under weak identification. Moreover, recall from Remark 3.3(iii) that the asymptotic bias of the OLS estimator is given by $b_{\hat{\beta}_{OLS}} = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho$. It follows that, by taking the ratio of the two, we obtain the relative bias measure

$$
\frac{b_{\hat{\beta}_{IV}}(\tau^2; k_{21})}{b_{\hat{\beta}_{OLS}}} = \left(\frac{1}{1+\tau^2}\right) - \frac{2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2. \tag{41}
$$

Note that the lead term of expression (41) is given by $(1+\tau^2)^{-1} = (1+\frac{\mu'\mu}{k_{21}})^{-1}$. When all available instruments are used so that IV = 2SLS, $(1+\frac{\mu'\mu}{k_{21}})^{-1} = (1+\frac{\mu'\mu}{k_2})^{-1}$ is simply the relative bias measure given in Staiger and Stock (1997), in the case where there is only a single endogenous explanatory variable. See section 3B and, in particular, section 6A of that paper. Hence, our result can be interpreted as giving a theoretical justification for the relative bias measure of Staiger and Stock (1997) for the case of only one endogenous regressor. Note that Staiger and Stock (1997) claims that this measure of relative bias arises from an approximation which holds for large k_{21} and/or large $\frac{\mu'\mu}{k_{21}}$. Our analysis, however, shows that an implicit assumption of holding $\frac{\mu'\mu}{k_{21}}$ fixed, rather than making $\frac{\mu'\mu}{k_{21}}$ large, as $k_{21} \to \infty$ is needed in order to extract such an approximation. Moreover, numerical calculations from Staiger and Stock (1997) show that the lead term $(1+\frac{\mu'\mu}{k_{21}})^{-1}$ gives a very good approximation for the relative bias of 2SLS to that of OLS, except in the cases with relatively small k_{21}. Hence, in the approximation (41), we include an extra correction term

$$
\frac{-2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right) \tag{42}
$$

to help give a better approximation when k_{21} is small. We will report the results of a numerical evaluation of our approximation in the next section.

(iii) Now, set

$$
\begin{aligned}
m_{\hat{\beta}_{IV}}(\tau^2; k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left(\frac{1}{1+\tau^2}\right)^2 \\
&\quad + \sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+\sigma_{uu}\sigma_{vv}^{-1}\rho^2\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left[1+3\left(\frac{1}{1+\tau^2}\right)\right. \\
&\left.+2\left(\frac{1}{1+\tau^2}\right)^2-6\left(\frac{1}{1+\tau^2}\right)^3\right]; \qquad (43)
\end{aligned}
$$

so that the notation $m_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ here denotes our approximation for the AMSE of the IV estimator under weak identification. Note that the lead term of $m_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ is given by $\sigma_{uu}\sigma_{vv}^{-1}\rho^2(1+\tau^2)^{-2}$, which is simply the square of the lead term of the bias approximation $b_{\hat{\beta}_{IV}}(\tau^2; k_{21})$: It follows that the variance component of the AMSE is of a lower order relative to the bias component in this expansion, so that the variance can be thought of as being negligible relative to the bias when the number of instruments is large relative to the value of τ^2, the population analogue of the first stage F-statistic.

(iv) Given the bias approximation (40) and the MSE approximation (43), we can, in fact, construct an approximation for the asymptotic variance of the IV estimator under weak identification as follows:

$$
\begin{aligned}
V_{\hat{\beta}_{IV}}(\tau^2; k_{21}) &= m_{\hat{\beta}_{IV}}(\tau^2; k_{21}) - (b_{\hat{\beta}_{IV}}(\tau^2; k_{21}))^2 \\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left(\frac{1}{1+\tau^2}\right)^2 \\
&\quad+\sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right) \\
&\quad+\sigma_{uu}\sigma_{vv}^{-1}\rho^2\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left[1+3\left(\frac{1}{1+\tau^2}\right)\right. \\
&\quad\left.+2\left(\frac{1}{1+\tau^2}\right)^2-6\left(\frac{1}{1+\tau^2}\right)^3\right] \\
&\quad-\sigma_{uu}\sigma_{vv}^{-1}\rho^2\left[\left(\frac{1}{1+\tau^2}\right)^2-\frac{4}{k_{21}}\left(\frac{1}{1+\tau^2}\right)^2+\right. \\
&\quad\left.+\frac{8}{k_{21}}\left(\frac{1}{1+\tau^2}\right)^3-\frac{4}{k_{21}}\left(\frac{1}{1+\tau^2}\right)^4\right]+O(k_{21}^{-2}) \\
&= \sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right) \\
&\quad+\sigma_{uu}\sigma_{vv}^{-1}\rho^2\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left[1+7\left(\frac{1}{1+\tau^2}\right)\right. \\
&\quad\left.-6\left(\frac{1}{1+\tau^2}\right)^2-2\left(\frac{1}{1+\tau^2}\right)^3\right]+O(k_{21}^{-2})
\end{aligned}
$$

$$= \widehat{V}_{\hat{\beta}_{IV}}(\tau^2; k_{21}) + O(k_{21}^{-2}); \text{ (say)}: \tag{44}$$

The following theorem shows that both our MSE approximation $\widehat{M}_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ and our variance approximation $\widehat{V}_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ are always non-negative, as one would expect variance and MSE to be.

Theorem 4.3: Given $\tau^2 \in [0; \infty)$;

(a) $\widehat{M}_{\hat{\beta}_{IV}}(\tau^2; k_{21}) \geq 0$ and

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

where $\widehat{M}_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ and $\widehat{V}_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ are defined by expressions (43) and (44), respectively.

(v) In part (e) of Theorem 3.1, we show that, for a fixed value of the concentration parameter $\mu'\mu$, the absolute magnitude of the bias $b_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ increases as the number of instruments used, i.e., k_{21}, increases. That result applies to the special case where the additional instruments are completely uncorrelated with the endogenous explanatory variable so that their inclusion leads to no increase in the value of $\mu'\mu$. In general, however, the use of additional instruments only which are only weakly correlated with the endogenous regressor will nevertheless lead to some increase in the value of the concentration parameter $\mu'\mu$: Hence, it would be of interest to study how the asymptotic bias and MSE functions vary with increases in the number of instruments used for the case where the inclusion of additional instruments increases the value of $\mu'\mu$ but not by enough to increase the value of the first stage F-statistic, or its population analogue τ^2. Below we give a theorem which characterizes the derivatives with respect to k_{21} of the asymptotic bias function and of our approximation of the AMSE function when τ^2.is held fixed.

Theorem 4.4: Let $b_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ be as defined in expression (9) and let $\widehat{M}_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ be as defined in expression (43). Then, it follows that:

(a)

$$\begin{aligned} \frac{\partial |b_{\hat{\beta}_{IV}}(\tau^2; k_{21})|}{\partial k_{21}} &= |\rho|\sigma_{uu}\sigma_{vv}^{-1} e^{-\frac{\tau^2 k_{21}}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\tau^2 k_{21}}{2}\right)^j}{j!} \frac{\gamma_j}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j)} \\ &> 0 \end{aligned} \tag{45}$$

and (b)

$$\frac{\partial \widehat{M}_{\hat{\beta}_{IV}}(\tau^2; k_{21})}{\partial k_{21}} = -\sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)\frac{1}{k_{21}^2}\left(\frac{1}{1+\tau^2}\right.$$

$$
-\sigma_{uu}\sigma_{vv}^{-1}\rho^2\frac{1}{k_{21}^2}\left(\frac{1}{1+\tau^2}\right)\left[1+3\left(\frac{1}{1+\tau^2}\right)+2\left(\frac{1}{1+\tau^2}\right)^2-6\left(\frac{1}{1+\tau^2}\right)^3\right]
$$
$$
< 0. \tag{46}
$$

From Theorem 4.4, we see that given a fixed value of τ^2 the absolute value of $b_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ is a monotonically increasing function of the number of instruments used k_{21} while $m_{\hat{\beta}_{IV}}(\tau^2; k_{21})$ is a monotonically decreasing function of k_{21}. This result suggests that there is a tradeoff between the bias and the variance of the IV estimator as k_{21} increases. It should be emphasized that several papers from the literature on the finite sample properties of the IV estimator under Gaussian assumptions (notably, Phillips (1980, 1983)) have already noted this tradeoff between the bias and variance, although as Buse (1992) pointed out the analysis of that literature focused exclusively on the case where the additional instruments are completely uncorrelated with the endogenous explanatory variable, i.e., the case where the concentration parameter $\mu'\mu$ is held fixed as k_{21} varied. Our result, on the other hand, shows that a tradeoff between the bias and MSE continues to hold even when the additional instruments are somewhat correlated with the endogenous regressor, as long as the correlation is not sufficient to increase the value of τ^2, the population analogue of the first stage F-statistic.

(vi) Note that given the near equivalence of the formulae for the asymptotic bias and MSE when instruments are weak in the sense of Staiger and Stock (1997) and the formulae for the exact bias and MSE under a fixed instrument/Gaussian model, as discussed in Remark 1(vi) of the last section, our results in this paper apply not only to the weak instrument case asymptotically but also to the exact finite sample case under Gaussian error assumptions. In particular, we note that the exact bias and MSE of the IV estimator under the assumption of a fixed instrument/Gaussian model can also be expanded in the manner given by expressions (38) and (39) of Theorem 4.1. It follows that the variance component of the exact MSE is also a term of a lower order in k_{21} relative to the bias component, so that when the number of instruments used is large, one can think of the variance as being negligible relative to the bias. Moreover, we note that for a fixed value of τ^2, the exact bias function increases while the exact MSE decreases with an increase in the number of instruments used, although this apparent tradeoff becomes less significant when as k_{21} becomes large.

5 Numerical Calculations

In this section, we summarize the results of a number of regressions used to examine the accuracy of our approximations of the IV bias and MSE functions. We tackle the issue of the usefulness of our approximations in three distinct manners. In our first set of numerical computations, we run regressions where the dependent variable is the IV bias (or MSE) calculated using the analytical formulae from above for certain values of k_{21}, $\mu'\mu$, and ρ; and the regresors are the functions of these three parameters which are used in the approximations discussed in the previous section. Note that based on the canonical model, the asymptotic OLS bias and ρ are related, as $\sigma_{vv} = 1$; $\sigma_{uu} = (1+\zeta^2)$; and $\rho = -\zeta/\sqrt{\sigma_{uu}}$; where ζ is the asymptotic OLS bias. Thus, $\zeta = -\sigma_{uu}^{1/2}\sigma_{uu}^{-1/2}\rho$. The grid of parameters for which values of the IV bias and MSE are calculated is: $k_{21} = \{3; 5; 7; 9; 11; \dots; 101\}$, $\mu'\mu = \{0; 2; 4; 6; 8; \dots; 100\}$, and $\zeta = \{-0.05; -0.1; -0.15; -0.2; -0.25; \dots; -1.0\}$. Thus, a total of 51000 unique observations are generated. Regression results are summarized in Tables 1 (bias) and 2 (MSE). Based on these findings, a number of clear conclusions emerge. First, our approximations based on holding $\mu'\mu/k_{21}$ fixed and letting $k_{21} \to \infty$ are very accurate for bias (see first 2 columns of entries in Table 1), relative bias (4th and 5th columns in Table 1), MSE (first 2 columns in Table 2) and relative MSE (4th and 5th columns in Table 1). Adjusted R^2 values for these regressions range from 0.9859 to 1.000. Second, note that each pair of regressions contains one set of regressors from the approximation which includes first and second order terms in the expansion, and one which includes only the first order term. While the fit of the regressions always improves when the second order terms are included as additional regressors, the improvement is small. For example, in column 1 of Table 1 note that our approximation yields an $\overline{R}^2$ value of 0.9998 when second order terms are included, and an $\overline{R}^2$ value of 0.9996 when only the first order term (and an intercept) is included. Thus, while the second order terms are generally useful, they add little to the approximations. Finally, note that the regression $\overline{R}^2$ values from the 3rd and 6th regression reported in each table are based on an alternative approximation in which ($\mu'\mu$ is fixed and $k_{21} \to \infty$). In particular ... (to be filled in with approximation formulae!!!!!!). Clearly, this approximation method is inferior based on our regression results, as evidenced by the substantially lower $\overline{R}^2$ values associated with this type of approximation.

In order to shed further light on the usefulness of our prefered approximation as well as the other approximation discussed above, we carried out a second set of computations. In particular, we compare the actual bias and MSE values used in the above regressions with approximate values calculated directly from the approximations given above (i.e. rather than estimating the coefficients associated with each regressor as was done in our first set of computations, we used the actual coefficients which are given in the formulae for the approximations). A summary of our findings based on these calculations is given in Figures 1 and 2, where actual bias (MSE) values are represented by solid lines, and the difference between the actual and approximate values for the two approximation methods are represented by small dash ($\mu'\mu/k_{21}$ fixed and $k_{21} \to \infty$ approximation) and large dash ($\mu'\mu$ fixed and $k_{21} \to \infty$): to be completed ...

Finally, in a third set of computations, we examined the impact that changing $\mu'\mu/k_{21}$ has on relative bias (MSE) when k_{21} is fixed (Figure 3), and the impact of changing k_{21} when $\mu'\mu/k_{21}$ is fixed (Figure 4). to be completed ....

6 Summary and Conclusions

We have formalized the notion of weak identification within the local-to-zero asymptotic framework of Staiger and Stock (1997), and derived explicit analytical formulae for the asymptotic bias and mean square error (MSE) of the IV estimator. In addition, we derive approximations for these formulae which are shown via a series of numerical computations to be very accurate. These results generalize earlier findings by Staiger and Stock (1997), and allow us to link systematically the earlier literature on IV estimation. For example, we show that our results apply not only to the weak instrument case asymptotically, but also to the finite sample case with fixed (possibly good) instruments and Gaussian errors, since our formulae correspond to the exact bias and MSE functionals when a fixed instrument/Gaussian model is assumed. We also show that in the special case where all available instruments are used and where the underlying simultaneous equations model has an orthonormal canonical structure, the bias and MSE formulae which we obtain are identical to the exact bias and MSE of the 2SLS estimator obtained by Richardson and Wu (1971) under Gaussian error assumptions. This result gives a partial confirmation to the Staiger-Stock assertion, based on intuitive arguments, that the limiting distribution of the 2SLS estimator derived under the more general assumptions of the Staiger-Stock local-to-zero asymptotic framework coincides with the exact distribution of the same estimator derived under the more restrictive assumptions of a fixed instrument/Gaussian model.

Appendix

Lemma A1: Let $\hat{\beta}_{IV,T}$ be the IV estimator given by expression (4) and suppose that (1), (2) and Assumptions 1 and 2 hold. Then,

$$\hat{\beta}_{IV,T} - \beta_0 \Rightarrow \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} v_1^{-1} v_2; \tag{47}$$

where

$$v_1 = (\lambda + Z_{v,1})'(\lambda + Z_{v,1}); \tag{48}$$

$$v_2 = (\lambda + Z_{v,1})' Z_{u,1}; \tag{49}$$

and

$$Z_{u,1} = \Omega_{11}^{-\frac{1}{2}\prime}(\Psi_{Z_1 u} - Q_{Z_1 X} Q_{XX}^{-1} \Psi_{X u})\sigma_{uu}^{-\frac{1}{2}}; \tag{50}$$

$$Z_{v,1} = \Omega_{11}^{-\frac{1}{2}\prime}(\Psi_{Z_1 v} - Q_{Z_1 X} Q_{XX}^{-1} \Psi_{X v})\sigma_{vv}^{-\frac{1}{2}}; \tag{51}$$

and where

$$\begin{pmatrix} Z_{u,1} \\ Z_{v,1} \end{pmatrix} \sim N\left(0; \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \otimes I_{k_{21}}\right): \tag{52}$$

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 $x \to \infty$

$$_1F_1(a; c; x) = \frac{\Gamma(c)}{\Gamma(a)} e^x x^{-(c-a)} \left[\sum_{j=0}^{\rho-1} \frac{(c-a)_j}{j!} x^{-j} + O(|x|^{-\rho})\right]: \tag{53}$$

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,

$$_1F_1(a; c; x) = e^x \left[\sum_{j=0}^{\rho-1} \frac{(c-a)_j(-x)^j}{(c)_j\, j!} x^{-j} + O(|c|^{-\rho})\right]: \tag{54}$$

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

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

$$
\begin{aligned}
E\left[\left(\chi^2_v(\mu'\mu)\right)^{-r}\right] &= 2^{-r}e^{-\frac{\mu'\mu}{2}}\sum_{j=0}^{\infty}\frac{(\frac{\mu'\mu}{2})^j}{j!}\frac{\Gamma\left(\frac{1}{2}(v+2j)-r\right)}{\Gamma\left(\frac{1}{2}(v+2j)\right)} \\
&= 2^{-r}e^{-\frac{\mu'\mu}{2}}\frac{(\frac{\mu'\mu}{2})^j}{j!}\frac{\Gamma\left(\frac{v}{2}-r\right)}{\Gamma(\frac{v}{2})}\,{}_1F_1\left(\frac{1}{2}v-r;\frac{1}{2}v;\frac{\mu'\mu}{2}\right). \qquad (55)
\end{aligned}
$$

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,

$$
E\left[\phi(w'w)w\right] = \theta E\left[\phi\left(\chi^2_{J+2}(\theta'\theta)\right)\right]. \qquad (56)
$$

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,

$$
E\left[\phi(w'w)ww'\right] = E\left[\phi\left(\chi^2_{J+2}(\theta'\theta)\right)\right]I_J + E\left[\phi\left(\chi^2_{J+4}(\theta'\theta)\right)\right]\theta\theta'. \qquad (57)
$$

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

Lemma A7: Let $\tau^2 \in [0, \infty)$; then, as $k_{21} \to \infty$;

(a)

$$
\begin{aligned}
&{}_1F_1\left(\frac{k_{21}}{2}-1;\frac{k_{21}}{2};\frac{\tau^2k_{21}}{2}\right)\exp\left\{-\frac{\tau^2k_{21}}{2}\right\} \\
= {} & \left(\frac{1}{1+\tau^2}\right) - \left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right) \\
& \times \left[2-4\left(\frac{1}{1+\tau^2}\right)+2\left(\frac{1}{1+\tau^2}\right)^2\right] \\
& + \left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^2\left[8+12\left(\frac{1}{1+\tau^2}\right)\right. \\
& \left. -32\left(\frac{1}{1+\tau^2}\right)^2+12\left(\frac{1}{1+\tau^2}\right)^3\right] + O(k_{21}^{-3}); \qquad (58)
\end{aligned}
$$

(b)

$$
\begin{aligned}
&{}_1F_1\left(\frac{k_{21}}{2}-2;\frac{k_{21}}{2}-1;\frac{\tau^2k_{21}}{2}\right)\exp\left\{-\frac{\tau^2k_{21}}{2}\right\} \\
&= \left(\frac{1}{1+\tau^2}\right) - \left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right) \\
&\quad\times\left[4-6\left(\frac{1}{1+\tau^2}\right)+2\left(\frac{1}{1+\tau^2}\right)^2\right] \\
&\quad+\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^2\left[12+20\left(\frac{1}{1+\tau^2}\right)\right. \\
&\quad\left.-44\left(\frac{1}{1+\tau^2}\right)^2+12\left(\frac{1}{1+\tau^2}\right)^3\right]+O(k_{21}^{-3}).
\end{aligned}
\tag{59}
$$

Proof: We shall prove only part (a) as the proof for part (b) follows in a similar 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

$$
\begin{aligned}
&{}_1F_1\left(\frac{k_{21}}{2}-1;\frac{k_{21}}{2};\frac{\tau^2k_{21}}{2}\right)\exp\left\{-\frac{\tau^2k_{21}}{2}\right\} \\
&= \frac{\Gamma\left(\frac{k_{21}}{2}\right)}{\Gamma\left(\frac{k_{21}}{2}-1\right)}\int_0^1\exp\left\{\frac{\tau^2k_{21}}{2}(t-1)\right\}t^{(k_{21}-4)/2}dt \\
&= \left(\frac{k_{21}-2}{2}\right)\int_0^1\exp\{k_{21}h_1(t)\}\,dt,
\end{aligned}
\tag{60}
$$

where

$$h_1(t)=\frac{\tau^2}{2}(t-1)+\frac{1}{2}\log t-\frac{2}{k_{21}}\log t. \tag{61}$$

Given the integral representation (60), we can obtain the expansion given by the right-hand side of expression (58) by applying a Laplace approximation to this integral representation. We note that the maximum of the integrand of (60) 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:

$$\left(\frac{k_{21}-2}{2}\right)\int_0^1\exp\{k_{21}h_1(t)\}\,dt = \left(\frac{k_{21}-2}{2}\right)\int_{1-1/\sqrt{k_{21}}}^1\exp\{k_{21}h_1(t)\}\,dt$$

$$+\left(\frac{k_{21}-2}{2}\right)\int_0^{1-1/\sqrt{k_{21}}}\exp\{k_{21}h_1(t)\}\,dt$$

$$= \quad I_1 + I_2 \quad (\text{say}); \qquad (62)$$

We shall handle I_2 first. Note that

$$\left(\frac{k_{21}-2}{2}\right)\int_0^{1-1/\sqrt{k_{21}}}\exp\{k_{21}h_1(t)\}\,dt$$

$$\le \left(\frac{k_{21}-2}{2}\right)\exp\left\{-\frac{\tau^2\sqrt{k_{21}}}{2}\right\}\left(1-\frac{1}{\sqrt{k_{21}}}\right)^{(k_{21}-2)/2}$$

$$= O\left(k_{21}\exp\left\{-\tau^2\sqrt{k_{21}}/2\right\}\left(1-1/\sqrt{k_{21}}\right)^{(k_{21}-2)/2}\right) \qquad (63)$$

Now, turning our attention to I_1, we first make the change of variable $v = t - 1$ and rewrite I_1 as

$$\left(\frac{k_{21}-2}{2}\right)\int_{-1/\sqrt{k_{21}}}^{0}\exp\{k_{21}h_2(v)\}\,dv \qquad (64)$$

where

$$h_2(v) = \frac{\tau^2}{2}v + \frac{1}{2}\log(1+v) - \frac{2}{k_{21}}\log(1+v). \qquad (65)$$

With this change of variable, we note that the maximum of the integrand in expression (64) in the interval $[-1/\sqrt{k_{21}}; 0]$ now occurs at the boundary point $v = 0$. To apply the Laplace approximation to expression (64), note that, by the Taylor theorem, we have, for $-1/\sqrt{k_{21}} \le v \le 0$; that

$$\left|h_2(v) - \sum_{i=1}^{6}\frac{h_2^{(i)}(0)}{i!}v^i\right| \le \frac{|v|^7}{7!}\sup_{u\in[-\frac{1}{\sqrt{k_{21}}};\frac{1}{\sqrt{k_{21}}}]}|h_2^{(7)}(u)|$$

$$\le \frac{\left(1/\sqrt{k_{21}}\right)^{\frac{7}{2}}}{7!}\left|\left(\frac{1}{2}-\frac{2}{\sqrt{k_{21}}}\right)6!\left(\frac{\sqrt{k_{21}}}{\sqrt{k_{21}}-1}\right)\right|$$

$$= \varepsilon(k_{21}) \quad (\text{say}); \qquad (66)$$

where $\varepsilon(k_{21}) = O(k_{21}^{-\frac{7}{2}})$. It follows that

$$\sum_{i=1}^{6}\frac{h_2^{(i)}(0)}{i!}v^i - \varepsilon(k_{21}) \le h_2(v) \le \sum_{i=1}^{6}\frac{h_2^{(i)}(0)}{i!}v^i + \varepsilon(k_{21}), \qquad (67)$$

so that

$$\left(\frac{k_{21}-2}{2}\right)\int_{-\frac{1}{\sqrt{k_{21}}}}^{0}\exp\left\{k_{21}\left(\sum_{i=1}^{6}\frac{h_2^{(i)}(0)}{i!}v^i - \varepsilon(k_{21})\right)\right\}dv$$

$$\leq \left(\frac{k_{21}-2}{2}\right)\int_{-1/\sqrt{k_{21}}}^{0} \exp\{k_{21}h_2(v)\}\,dv$$

$$\leq \left(\frac{k_{21}-2}{2}\right)\int_{-\frac{1}{\sqrt{k_{21}}}}^{0} \exp\left\{k_{21}\left(\sum_{i=1}^{6}\frac{h_2^{(i)}(0)}{i!}v^i + \varepsilon(k_{21})\right)\right\}dv; \tag{68}$$

where the derivatives of $h_2(v)$ evaluated at $v = 0$ have the explicit forms:

$$h_2'(0) = \left(\frac{1+\tau^2}{2}\right) - \frac{2}{\sqrt{k_{21}}} \tag{69}$$

and

$$h_2^{(i)}(0) = (-1)^{i-1}(i-1)!\left[\frac{1}{2} - \frac{2}{k_{21}}\right] \quad \text{for } i = 2, \ldots, 6. \tag{70}$$

Let I_3 denotes the upper bound integral in expression (68). To evaluate I_3, we rewrite it as

$$\begin{aligned} I_3 &= \left(\frac{k_{21}-2}{2}\right)\int_{-\frac{1}{\sqrt{k_{21}}}}^{0} \exp\left\{k_{21}h_2'(0)v\right\} \\ &\quad \exp\left\{k_{21}\left(\sum_{i=2}^{6}\frac{h_2^{(i)}(0)}{i!}v^i\right)\right\}\exp\{k_{21}\varepsilon(k_{21})\}\,dv. \end{aligned} \tag{71}$$

Expanding the latter two exponentials in the integrand above in power series and integrating term-by-term, we obtain, after some tedious but straightforward calculations, the approximation

$$\begin{aligned} I_3 &= \left(\frac{k_{21}-2}{2}\right)\left[2\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right) + 8\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^2 \right. \\ &\quad +32\left(\frac{1}{k_{21}}\right)^3\left(\frac{1}{1+\tau^2}\right)^3 - 4\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^3 \\ &\quad +16\left(\frac{1}{k_{21}}\right)^3\left(\frac{1}{1+\tau^2}\right)^3 - 48\left(\frac{1}{k_{21}}\right)^3\left(\frac{1}{1+\tau^2}\right)^4 \\ &\quad \left. -16\left(\frac{1}{k_{21}}\right)^3\left(\frac{1}{1+\tau^2}\right)^4 + 24\left(\frac{1}{k_{21}}\right)^3\left(\frac{1}{1+\tau^2}\right)^5\right] + O(k_{21}^{-\frac{7}{2}}) \\ &= \left(\frac{1}{1+\tau^2}\right) - \left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\left[2 - 4\left(\frac{1}{1+\tau^2}\right) \right. \\ &\quad \left. + 2\left(\frac{1}{1+\tau^2}\right)^2\right] + \left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^2\left[8 + 12\left(\frac{1}{1+\tau^2}\right) \right. \\ &\quad \left. -32\left(\frac{1}{1+\tau^2}\right)^2 + 12\left(\frac{1}{1+\tau^2}\right)^3\right] + O(k_{21}^{-3}). \end{aligned} \tag{72}$$

By a similar argument, it can be shown that the lower bound integral in expression (68) can also be approximated by the right-hand side of expression (72). It follows that

$$
\begin{aligned}
&\left(\frac{k_{21}-2}{2}\right)\int_{-1/\sqrt{k_{21}}}^{0} \exp\{k_{21}h_2(v)\}\,dv \\
= &\left(\frac{1}{1+\tau^2}\right) - \left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\left[2 - 4\left(\frac{1}{1+\tau^2}\right)\right. \\
&\left. + 2\left(\frac{1}{1+\tau^2}\right)^2\right] + \left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)\left[8 + 12\left(\frac{1}{1+\tau^2}\right)\right. \\
&\left. - 32\left(\frac{1}{1+\tau^2}\right)^2 + 12\left(\frac{1}{1+\tau^2}\right)^3\right] + O(k_{21}^{-3}).
\end{aligned}
\tag{73}
$$

Finally, the result given in part (a) follows immediately from expressions (63) and (73). □

Proof of Theorem 3.1: To show part (a), we note that by Lemma A1

$$W_T = \hat{\beta}_{IV,T} - \beta_0 \Rightarrow \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} v_1^{-1} v_2 \equiv w_1 \text{ (say)}. \tag{74}$$

Moreover, given Assumption 3, we have by Theorem 5.4 of Billingsley (1968) that

$$
\begin{aligned}
\lim_{T\to\infty} E(W_T) &= \lim_{T\to\infty} E\left[\hat{\beta}_{IV,T} - \beta_0\right] \\
&= E\left[\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} v_1^{-1} v_2\right] \\
&= E(W).
\end{aligned}
\tag{75}
$$

It follows that to derive the asymptotic bias of $\hat{\beta}_{IV}$, 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 (52), we can write

$$Z_{u,1} = Z_{v,1}\rho + Z_{u1.v1}, \tag{76}$$

where $Z_{u1.v1} \sim N(0,(1-\rho^2)I_{k_{21}})$ represents the projection error and is, thus, independent of Z_v. Next, we rewrite the limiting random variable W as

$$
\begin{aligned}
W &= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} v_1^{-1} v_2 \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\left[(\mu + Z_{v,1})'(\mu + Z_{v,1})\right]^{-1}(\mu + Z_{v,1})'Z_{u,1} \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\left[(\mu + Z_{v,1})'(\mu + Z_{v,1})\right]^{-1}(\mu + Z_{v,1})'(Z_{v,1}\rho + Z_{u1.v1}),
\end{aligned}
\tag{77}
$$

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

$$
\begin{aligned}
E(W) &= E_{Z_{v,1}}\left[E_{Z_{u,1}|Z_{v,1}}\left(\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'(Z_{v,1}\rho+Z_{u1\cdot v1})\right)\right] \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}E_{Z_v}\left[\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'Z_{v,1}\rho\right]; \qquad (78)
\end{aligned}
$$

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 (78), we note that

$$
\begin{aligned}
&\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}E_{Z_{v,1}}\left[\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'Z_{v,1}\rho\right] \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}E_{Z_{v,1}}\left[\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}(\mu+Z_{v,1})'(\mu+Z_{v,1}-\mu)\rho\right] \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho-\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}E_{Z_{v,1}}\left[\left[(\mu+Z_{v,1})'(\mu+Z_{v,1})\right]^{-1}\mu\rho\right]. \qquad (79)
\end{aligned}
$$

Note further that

$$(\mu+Z_{v,1})\sim N(\mu; I_{k_{21}})$$

so that $(\mu+Z_{v,1})'(\mu+Z_{v,1})\sim\chi^2_{k_{21}}(\mu'\mu)$. It follows from Lemma A5 that

$$\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho-\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\mu'\mu E\left[\frac{1}{\chi^2_{k_{21}+2}(\mu'\mu)}\right]. \qquad (80)$$

Finally, applying Lemma A4 to (80), we obtain

$$
\begin{aligned}
\lim_{T\to\infty}E\left[\widehat{\beta}_{IV,T}-\beta_0\right] &= E(W) \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\times \\
&\quad\left[1-\left(\frac{\mu'\mu}{2}\right)e^{-\frac{\mu'\mu}{2}}\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{\Gamma\left(\frac{k_{21}}{2}+j\right)}{\Gamma\left(\frac{k_{21}}{2}+j+1\right)}\right] \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\times \\
&\quad\left[1-\left(\frac{\mu'\mu}{2}\right)e^{-\frac{\mu'\mu}{2}}\frac{\Gamma\left(\frac{k_{21}}{2}\right)}{\Gamma\left(\frac{k_{21}}{2}+1\right)}{}_1F_1\left(\frac{k_{21}}{2};\frac{k_{21}}{2}+1;\frac{\mu'\mu}{2}\right)\right]; \qquad (81)
\end{aligned}
$$

Now, applying the recurrence relation

$$
\begin{aligned}
z\,{}_1F_1(\alpha+1;\gamma+1;z) &= \gamma\left[{}_1F_1(\alpha+1;\gamma;z)\right. \\
&\quad\left.-\,{}_1F_1(\alpha;\gamma;z)\right] \qquad (82)
\end{aligned}
$$

and noting that ${}_1F_1(\alpha;\alpha;z) = e^z$, we obtain the desired result.

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

$$
\begin{aligned}
&\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1 - \left(\frac{\mu'\mu}{2}\right) e^{-\frac{\mu'\mu}{2}} \frac{\Gamma(\frac{k_{21}}{2})}{\Gamma(\frac{k_{21}}{2}+1)}\, {}_1F_1\left(\frac{k_{21}}{2};\frac{k_{21}}{2}+1;\frac{\mu'\mu}{2}\right)\right] \\
= \;& \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1 - \left(\frac{\mu'\mu}{2}\right) e^{-\frac{\mu'\mu}{2}} \frac{\Gamma(\frac{k_{21}}{2})}{\Gamma(\frac{k_{21}}{2}+1)} \frac{\Gamma(\frac{k_{21}}{2}+1)}{\Gamma(\frac{k_{21}}{2})} e^{\frac{\mu'\mu}{2}}\left(\frac{\mu'\mu}{2}\right)^{-1}\left(1 + O(|\mu'\mu|^{-1})\right)\right] \\
= \;& \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1 - \left(1 + O(|\mu'\mu|^{-1})\right)\right] \\
= \;& O(|\mu'\mu|^{-1}); \qquad (83)
\end{aligned}
$$

which proves (b).

To show part (c), note that

$$
\lim_{k_{21}\to\infty} \frac{(\frac{1}{2}k_{21} + 1 - \frac{1}{2}k_{21})\left(\frac{\mu'\mu}{2}\right)}{(\frac{1}{2}k_{21}+1)} = \lim_{k_2\to\infty}\left(\frac{\mu'\mu}{k_{21}+2}\right) = 0: \qquad (84)
$$

Hence, direct application of Lemma A3 gives

$$
\begin{aligned}
&\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1 - \left(\frac{\mu'\mu}{2}\right) e^{-\frac{\mu'\mu}{2}} \frac{\Gamma(\frac{k_{21}}{2})}{\Gamma(\frac{k_{21}}{2}+1)}\, {}_1F_1\left(\frac{k_{21}}{2};\frac{k_{21}}{2}+1;\frac{\mu'\mu}{2}\right)\right] \\
= \;& \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1 - \left(\frac{\mu'\mu}{2}\right) e^{-\frac{\mu'\mu}{2}} \frac{\Gamma(\frac{k_{21}}{2})}{\Gamma(\frac{k_{21}}{2}+1)} e^{\frac{\mu'\mu}{2}}\left(1 + O(|k_{21}|^{-1})\right)\right] \\
= \;& \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1 - \left(\frac{\mu'\mu}{2}\right)\left(\frac{1}{\frac{k_{21}}{2}}\right)\left(1 + O(|k_{21}|^{-1})\right)\right] \\
= \;& \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[\left(1 + O(|k_{21}|^{-1})\right)\right] \\
\to \;& \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho \quad \text{as} \quad k_{21} \to \infty; \qquad (85)
\end{aligned}
$$

which proves (c)

To show (d), note that from expression (81) above that

$$
\begin{aligned}
b_{IV}(\mu'\mu;k_{21}) &= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1 - \left(\frac{\mu'\mu}{2}\right) e^{-\frac{\mu'\mu}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!} \frac{\Gamma(\frac{k_{21}}{2}+j)}{\Gamma(\frac{k_{21}}{2}+j+1)}\right] \\
&= \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left[1 - e^{-\frac{\mu'\mu}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\mu'\mu}{2}\right)^{j+1}}{j!} \frac{1}{(\frac{k_{21}}{2}+j)}\right] \qquad (86)
\end{aligned}
$$

Let

$$f(\mu'\mu; k_{21}) = \left[1 - e^{-\frac{\mu'\mu}{2}} \sum_{j=0} \frac{\left(\frac{\mu'\mu}{2}\right)^{j+1}}{j!} \frac{1}{(\frac{k_{21}}{2} + j)}\right] \tag{87}$$

and note further that

$$f(\mu'\mu = 0; k_{21}) = 1 \tag{88}$$

and from the proof of part (b), we know that

$$\lim_{\mu'\mu \to \infty} f(\mu'\mu; k_{21}) = 0 \tag{89}$$

Moreover, note that, under the assumption that $k_{21} \geq 4$;

$$\begin{aligned}
\frac{\partial f(\mu'\mu; k_{21})}{\partial(\mu'\mu)} &= (\frac{1}{2})e^{-\frac{\mu'\mu}{2}} \left[\sum_{j=0} \frac{\left(\frac{\mu'\mu}{2}\right)^{j+1}}{j!} \frac{1}{(\frac{k_{21}}{2} + j)} - \sum_{j=0} \frac{(j+1)\left(\frac{\mu'\mu}{2}\right)^{j}}{j!} \frac{1}{(\frac{k_{21}}{2} + j)}\right] \\
&= (\frac{1}{2})e^{-\frac{\mu'\mu}{2}} \left[\sum_{i=1} \frac{\left(\frac{\mu'\mu}{2}\right)^{i}}{(i-1)!} \frac{1}{(\frac{k_{21}}{2} + i - 1)} - \sum_{j=0} \frac{(j+1)\left(\frac{\mu'\mu}{2}\right)^{j}}{j!} \frac{1}{(\frac{k_{21}}{2} + j)}\right] \\
&= (\frac{1}{2})e^{-\frac{\mu'\mu}{2}} \left[\sum_{i=0} \frac{i\left(\frac{\mu'\mu}{2}\right)^{i}}{i!} \frac{(\frac{k_{21}}{2} + i)}{(\frac{k_{21}}{2} + i - 1)(\frac{k_{21}}{2} + i)} \right. \\
&\quad - \sum_{j=0} \frac{j\left(\frac{\mu'\mu}{2}\right)^{j}}{j!} \frac{(\frac{k_{21}}{2} + j - 1)}{(\frac{k_{21}}{2} + j - 1)(\frac{k_{21}}{2} + j)} \\
&\quad \left. - \sum_{j=0} \frac{\left(\frac{\mu'\mu}{2}\right)^{j}}{j!} \frac{(\frac{k_{21}}{2} + j - 1)}{(\frac{k_{21}}{2} + j - 1)(\frac{k_{21}}{2} + j)}\right] \\
&= (\frac{1}{2})e^{-\frac{\mu'\mu}{2}} \left[\sum_{j=0} \frac{j\left(\frac{\mu'\mu}{2}\right)^{j}}{j!} \frac{1}{(\frac{k_{21}}{2} + j - 1)(\frac{1}{2}k_{21} + j)} \right. \\
&\quad \left. - \sum_{j=0} \frac{\left(\frac{\mu'\mu}{2}\right)^{j}}{j!} \frac{(\frac{k_{21}}{2} + j - 1)}{(\frac{k_{21}}{2} + j - 1)(\frac{k_{21}}{2} + j)}\right] \\
&= (-\frac{1}{2})\left[e^{-\frac{\mu'\mu}{2}} \sum_{j=0} \frac{\left(\frac{\mu'\mu}{2}\right)^{j}}{j!} \frac{(\frac{k_{21}}{2} - 1)}{(\frac{k_{21}}{2} + j - 1)(\frac{1}{2}k_{21} + j)}\right] < 0;
\end{aligned} \tag{90}$$

where term-by-term differentiation is justified by the absolute convergence of the infinite series in (81). It follows from (88), (89), and (90) that

$$0 \le f(\mu'\mu; k_{21}) \le 1 \tag{91}$$

and is a monotonically decreasing function of $(\mu'\mu)$ for $(\mu'\mu) \in [0, \infty)$. Moreover, from (86), we have that

$$|b_{IV}(\mu'\mu; k_{21})| = |\rho|\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} f(\mu'\mu; k_{21}) \tag{92}$$

so that $|b_{IV}(\mu'\mu; k_{21})|$ depends on $\mu'\mu$ only through the factor $f(\mu'\mu; k_{21})$. Hence, it follows also that $|b_{IV}(\mu'\mu; k_{21})|$ is a monotonically decreasing function of $\mu'\mu$ for $\mu'\mu \in [0, \infty)$.

To show (e), note that due to the absolute convergence of the infinite series in (81), we can differentiate term-by-term to obtain

$$\frac{\partial f(\mu'\mu; k_{21})}{\partial k_{21}} = e^{-\frac{\mu'\mu}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\mu'\mu}{2}\right)^{j+1}}{j!} \frac{\left(\frac{1}{2}\right)}{\left(\frac{k_{21}}{2} + j\right)^2} > 0. \tag{93}$$

Hence, $f(\mu'\mu; k_{21})$ is a monotonically increasing function of k_{21}. It then follows from (92) that $|b_{IV}(\mu'\mu; k_{21})|$ is a monotonically increasing function of k_{21}. □

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

$$\begin{aligned} \lim_{T\to\infty} E\left[\hat{\beta}_{IV,T} - \beta_0\right]^2 &= \lim_{T\to\infty} E(W_T^2) \\ &= E(W^2) \\ &= E[\sigma_{uu}\sigma_{vv}^{-1} v_1^{-1} v_2^2 v_1^{-1}]. \end{aligned} \tag{94}$$

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 $E[\sigma_{uu}\sigma_{vv}^{-1} v_1^{-1} v_2^2 v_1^{-1}]$.

To proceed, note by expressions (52) and (75) that we can write

$$\begin{aligned} W^2 &= \sigma_{uu}\sigma_{vv}^{-1} v_1^{-1} v_2^2 v_1^{-1} \\ &= \sigma_{uu}\sigma_{vv}^{-1}\left[(\mu + Z_{v;1})'(\mu + Z_{v;1})\right]^{-1}(\mu + Z_{v;1})'(Z_{v;1}\rho + Z_{u1\cdot v1})(Z_{v;1}\rho + Z_{u1\cdot v1})' \\ &\quad (\mu + Z_{v;1})\left[(\mu + Z_{v;1})'(\mu + Z_{v;1})\right]^{-1}; \end{aligned} \tag{95}$$

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

$$\begin{aligned}
E(W^2) &= \sigma_{uu}\sigma_{vv}^{-1}\Big[E_{Z_{v;1}}\Big[E_{Z_{u;1}|Z_{v;1}}\Big(\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}(\mu+Z_{v;1})'(Z_{v;1}\rho+Z_{u1\cdot v1})(Z_{v;1}\rho+Z_{u1\cdot v1})' \\
&\qquad (\mu+Z_{v;1})\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}\Big)\Big]\Big] \\
&= \sigma_{uu}\sigma_{vv}^{-1}\Big[E_{Z_{v;1}}\Big(\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}(\mu+Z_{v;1})'\left[Z_{v;1}Z_{v;1}'\rho+(1-\rho^2)I_{k_{21}}\right) \\
&\qquad (\mu+Z_{v;1})\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}\Big)\Big] \\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 E_{Z_{v;1}}\Big(\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}(\mu+Z_{v;1})'Z_{v;1}Z_{v;1}'(\mu+Z_{v;1}) \\
&\qquad \left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}\Big) \\
&\qquad +\sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)E_{Z_{v;1}}\Big(\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}\Big) \\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 E_{Z_{v;1}}\Big(\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}(\mu+Z_{v;1})'(\mu+Z_{v;1}-\mu) \\
&\qquad (\mu+Z_{v;1}-\mu)'(\mu+Z_{v;1})\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}\Big) \\
&\qquad +\sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)E_{Z_{v;1}}\Big(\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}\Big) \\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\Big\{ 1-2E_{Z_{v;1}}\Big(\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}(\mu+Z_{v;1})'\mu\Big) \\
&\qquad +E_{Z_{v;1}}\Big(\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}\mu'(\mu+Z_{v;1})(\mu+Z_{v;1})'\mu \\
&\qquad \left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}\Big) \\
&\qquad +\left(\frac{1-\rho^2}{\rho^2}\right)E_{Z_{v;1}}\Big(\left[(\mu+Z_{v;1})'(\mu+Z_{v;1})\right]^{-1}\Big)\Big\}; \qquad (96)
\end{aligned}$$

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 $(\mu+Z_{v;1}) \sim N(\lambda; I_{k_{21}})$ and $(\mu+Z_{v;1})'(\mu+Z_{v;1}) \sim \chi^2_{k_{21}}(\mu'\mu)$; we can apply Lemma A5 and A6 to (96) above to obtain

$$\begin{aligned}
E(W^2) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\Bigg\{ 1-2E\left(\frac{1}{\chi^2_{k_{21}}(\mu'\mu)}\right)\mu'\mu \\
&\qquad \mu'\left[E\left(\frac{1}{\chi^2_{k_{21}+2}(\mu'\mu)}\right)^2 I_{k_{21}}+E\left(\frac{1}{\chi^2_{k_{21}+4}(\mu'\mu)}\right)^2\mu\mu'\right]\mu \\
&\qquad +\left(\frac{1-\rho^2}{\rho^2}\right)E\left(\frac{1}{\chi^2_{k_{21}}(\mu'\mu)}\right)\Bigg\}. \qquad (97)
\end{aligned}$$

Finally, applying Lemma A4 to expression (97) above, we obtain

$$
\begin{aligned}
\lim_{T \to \infty} E\left[\hat{\beta}_{IV,T} - \beta_0\right]^2 &= E(W^2) \\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2 \times \\
&\quad \left\{ 1 - (\mu'\mu) e^{-\frac{\mu'\mu}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!} \frac{\Gamma(\frac{k_{21}}{2}+j)}{\Gamma(\frac{k_{21}}{2}+j+1)} \right. \\
&\quad + \left(\frac{\mu'\mu}{4}\right) e^{-\frac{\mu'\mu}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!} \frac{\Gamma(\frac{k_{21}}{2}+j-1)}{\Gamma(\frac{k_{21}}{2}+j+1)} \\
&\quad + \left(\frac{\mu'\mu}{2}\right)^2 e^{-\frac{\mu'\mu}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!} \frac{\Gamma(\frac{k_{21}}{2}+j)}{\Gamma(\frac{k_{21}}{2}+j+2)} \\
&\quad \left. + \left(\frac{1-\rho^2}{2\rho^2}\right) e^{-\frac{\mu'\mu}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!} \frac{\Gamma(\frac{k_{21}}{2}+j-1)}{\Gamma(\frac{k_{21}}{2}+j)} \right\} \\
&= 1 - \left(\mu'\mu\right) e^{-\frac{\mu'\mu}{2}} \frac{\Gamma(\frac{k_{21}}{2})}{\Gamma(\frac{k_{21}}{2}+1)} {}_1F_1\left(\frac{k_{21}}{2}; \frac{k_{21}}{2}+1; \frac{\mu'\mu}{2}\right) \\
&\quad + \left(\frac{\mu'\mu}{4}\right) e^{-\frac{\mu'\mu}{2}} \frac{\Gamma(\frac{k_{21}}{2}-1)}{\Gamma(\frac{k_{21}}{2}+1)} {}_1F_1\left(\frac{k_{21}}{2}-1; \frac{k_{21}}{2}+1; \frac{\mu'\mu}{2}\right) \\
&\quad + \left(\frac{\mu'\mu}{2}\right)^2 e^{-\frac{\mu'\mu}{2}} \frac{\Gamma(\frac{k_{21}}{2})}{\Gamma(\frac{k_{21}}{2}+2)} {}_1F_1\left(\frac{k_{21}}{2}; \frac{k_{21}}{2}+2; \frac{\mu'\mu}{2}\right) \\
&\quad + \left(\frac{1-\rho^2}{2\rho^2}\right) e^{-\frac{\mu'\mu}{2}} \frac{\Gamma(\frac{k_{21}}{2}-1)}{\Gamma(\frac{k_{21}}{2})} {}_1F_1\left(\frac{k_{21}}{2}-1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right).
\end{aligned}
\tag{98}
$$

By successive application of the recurrence relation

$$
\begin{aligned}
z\, {}_1F_1(\alpha+1; \gamma+1; z) &= \gamma\, [{}_1F_1(\alpha+1; \gamma; z) \\
&\quad - {}_1F_1(\alpha; \gamma; z)]
\end{aligned}
\tag{99}
$$

noting that ${}_1F_1(\alpha; \alpha; z) = e^z$, we obtain the desired result.

To show (b), note that $\frac{\mu'\mu}{2} > 0$; $\frac{k_{21}}{2} - 1 > 0$; $\frac{k_{21}}{2} > 0$; $\frac{k_{21}}{2} + 1 > 0$; and $\frac{k_{21}}{2} + 2 > 0$; for $k_{21} \geq 4$; and, hence, we can apply Lemma A2 to each of the confluent hypergeometric functions ${}_1F_1(\cdot;\cdot;\cdot)$ which appear in (98) to obtain

$$
m_{\hat{\beta}_{IV}}(\mu'\mu; k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 + \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}} \times
$$

$$
\begin{aligned}
& \Bigg[(\tfrac{1}{2}) \left(\frac{\mu'\mu}{2}\right)^2 e^{-\frac{\mu'\mu}{2}} \left(\frac{\mu'\mu}{2}\right)^{-2} \left[1 + O(|\mu'\mu|^{-1})\right] \\
& + \left(\frac{\mu'\mu}{2}\right)^2 e^{-\frac{\mu'\mu}{2}} \left(\frac{\mu'\mu}{2}\right)^{-2} \left[1 + O(|\mu'\mu|^{-1})\right] \\
& + \left(\frac{1-\rho^2}{2\rho^2}\right)^2 e^{-\frac{\mu'\mu}{2}} \left(\frac{\mu'\mu}{2}\right)^{-1} \left[1 + O(|\mu'\mu|^{-1})\right] \\
& - 2 \left(\frac{\mu'\mu}{2}\right)^2 e^{-\frac{\mu'\mu}{2}} \left(\frac{\mu'\mu}{2}\right)^{-1} \left[1 + O(|\mu'\mu|^{-1})\right] \Bigg] \\
= \; & \sigma_{uu}\sigma_{VV}^{-1}\rho^2 + \sigma_{uu}\sigma_{VV}^{-1}\rho^2 \left[-1 \left[1 + O(|\mu'\mu|^{-1})\right] \right] \\
= \; & O(|\mu'\mu|^{-1}); \qquad (100)
\end{aligned}
$$

which proves (b).

To show (c), note that each of the ${}_1F_1(\cdot;\cdot;\cdot)$ functions appearing in expression (98) satisfies the conditions of Lemma A3 so we may apply this Lemma to obtain

$$
\begin{aligned}
m_{\hat{b}_{IV}}(\mu'\mu; k_{21}) = \; & \sigma_{uu}\sigma_{VV}^{-1}\rho^2 + \sigma_{uu}\sigma_{VV}^{-1}\rho^2 \times \\
& \Bigg[\left(\frac{\mu'\mu}{4}\right) e^{-\frac{\mu'\mu}{2}} \frac{1}{\left(\frac{k_{21}}{2}\right)\left(\frac{k_{21}}{2} - 1\right)} e^{\frac{\mu'\mu}{2}} \left[1 + O(|k_{21}|^{-1})\right] \\
& + \left(\frac{\mu'\mu}{2}\right)^2 e^{-\frac{\mu'\mu}{2}} \frac{1}{\left(\frac{k_{21}}{2} + 1\right)\left(\frac{k_{21}}{2}\right)} e^{\frac{\mu'\mu}{2}} \left[1 + O(|k_{21}|^{-1})\right] \\
& + \left(\frac{1-\rho^2}{2\rho^2}\right) \frac{1}{\left(\frac{k_{21}}{2} - 1\right)} e^{\frac{\mu'\mu}{2}} \left[1 + O(|k_{21}|^{-1})\right] \\
& - (\mu'\mu) e^{-\frac{\mu'\mu}{2}} \frac{1}{\left(\frac{k_{21}}{2}\right)} e^{\frac{\mu'\mu}{2}} \left[1 + O(|k_{21}|^{-1})\right] \Bigg] \\
= \; & \sigma_{uu}\sigma_{VV}^{-1}\rho^2 + O(|k_{21}|^{-1}) \\
\to \; & \sigma_{uu}\sigma_{VV}^{-1}\rho^2 \quad \text{as } k_{21} \to \infty; \qquad (101)
\end{aligned}
$$

which proves (c).

To show part (d), it suffices to show that

$$
\frac{\partial m_{\hat{b}_{2SLS}}(\mu'\mu; k_{21})}{\partial(\mu'\mu)} < 0
$$

for all fixed integer $k_{21} \geq 4$: To proceed, write (98) as

$$
m_{\hat{b}_{IV}}(\mu'\mu; k_{21}) = \sigma_{uu}\sigma_{VV}^{-1}\rho^2 + \sigma_{uu}\sigma_{VV}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}} \times
$$

$$\left[\left(\frac{1}{2}\right)\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^{j+1}}{j!}\frac{1}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)}\right.$$
$$+\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^{j+2}}{j!}\frac{1}{(\frac{k_{21}}{2}+j+1)(\frac{k_{21}}{2}+j)}$$
$$+\left(\frac{1-\rho^2}{2\rho^2}\right)\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^{j}}{j!}\frac{1}{(\frac{k_{21}}{2}+j-1)}$$
$$\left.-2\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^{j+1}}{j!}\frac{1}{(\frac{k_{21}}{2}+j)}\right] \qquad (102)$$

Now, noting that term-by-term differentiation of (102) is allowed due to the absolute convergence of the series involved, we have

$$\frac{\partial m_{\hat{b}_{IV}}(\mu'\mu;k_{21})}{\partial(\mu'\mu)} = \left(\frac{1}{2}\right)\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}} \times$$
$$\left\{\left[\left(\frac{1}{2}\right)\sum_{j=0}^{\infty}\frac{(j+1)\left(\frac{\mu'\mu}{2}\right)^{j}}{j!}\frac{1}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)}\right.\right.$$
$$\left.-\left(\frac{1}{2}\right)\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^{j+1}}{j!}\frac{1}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)}\right]$$
$$+\left[\sum_{j=0}^{\infty}\frac{(j+2)\left(\frac{\mu'\mu}{2}\right)^{j+1}}{j!}\frac{1}{(\frac{k_{21}}{2}+j+1)(\frac{k_{21}}{2}+j)}\right.$$
$$\left.-\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^{j+2}}{j!}\frac{1}{(\frac{k_{21}}{2}+j+1)(\frac{k_{21}}{2}+j)}\right]$$
$$+\left[\left(\frac{1-\rho^2}{2\rho^2}\right)\sum_{j=0}^{\infty}\frac{j\left(\frac{\mu'\mu}{2}\right)^{j-1}}{j!}\frac{1}{(\frac{k_{21}}{2}+j-1)}\right.$$
$$\left.-\left(\frac{1-\rho^2}{2\rho^2}\right)\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^{j}}{j!}\frac{1}{(\frac{k_{21}}{2}+j-1)}\right]$$
$$-\left[2\sum_{j=0}^{\infty}\frac{(j+1)\left(\frac{\mu'\mu}{2}\right)^{j}}{j!}\frac{1}{(\frac{k_{21}}{2}+j)}\right.$$

$$
-2\sum_{j=0}^{\infty}\frac{\left(\frac{1'1}{2}\right)^{j+1}}{j!}\frac{1}{(\frac{k_{21}}{2}+j)}\Bigg]\Bigg\}
$$

$$
= \left(\frac{1}{2}\right)\sigma_{uu}\sigma_{vv}^{-1}\rho^{2}e^{-\frac{1'1}{2}}\{A+B+C+D\};\ \text{(say)}, \tag{103}
$$

where A, B, C, and D are the expressions inside the square brackets. Focusing on A and letting $i = j + 1$, we have that

$$
\begin{aligned}
A &= \left(\frac{1}{2}\right)\left[\sum_{j=0}^{\infty}\frac{(j+1)\left(\frac{1'1}{2}\right)^{j}}{j!}\frac{1}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)}\right. \\
&\quad \left. -\sum_{i=1}^{\infty}\frac{\left(\frac{1'1}{2}\right)^{i}}{(i-1)!}\frac{1}{(\frac{k_{21}}{2}+i-1)(\frac{k_{21}}{2}+i-2)}\right] \\
&= \left(\frac{1}{2}\right)\left[\sum_{j=0}^{\infty}\frac{(j+1)\left(\frac{1'1}{2}\right)^{j}}{j!}\frac{(\frac{k_{21}}{2}+j-2)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}\right. \\
&\quad \left. -\sum_{i=1}^{\infty}\frac{i\left(\frac{1'1}{2}\right)^{i}}{i!}\frac{(\frac{k_{21}}{2}+i)}{(\frac{k_{21}}{2}+i)(\frac{k_{21}}{2}+i-1)(\frac{k_{21}}{2}+i-2)}\right] \\
&= \left(\frac{1}{2}\right)\left[\sum_{j=0}^{\infty}\frac{\left(\frac{1'1}{2}\right)^{j}}{j!}\frac{(\frac{k_{21}}{2}+j-2)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}\right. \\
&\quad \left. -\sum_{j=0}^{\infty}\frac{2j\left(\frac{1'1}{2}\right)^{j}}{j!}\frac{1}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}\right] \\
&= \left(\frac{1}{2}\right)\left[\sum_{j=0}^{\infty}\frac{\left(\frac{1'1}{2}\right)^{j}}{j!}\frac{(\frac{k_{21}}{2}+j-2)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}\right].
\end{aligned} \tag{104}
$$

Following similar calculations, we see that

$$
B = \sum_{j=0}^{\infty}\frac{j\left(\frac{1'1}{2}\right)^{j}}{j!}\frac{(k_{21}-2)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}, \tag{105}
$$

$$
C = -\left(\frac{1-\rho^{2}}{2\rho^{2}}\right)\sum_{j=0}^{\infty}\frac{\left(\frac{1'1}{2}\right)^{j}}{j!}\frac{1}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)}, \tag{106}
$$

$$D = 2\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{\left(\frac{k_{21}}{2}-1\right)}{\left(\frac{k_{21}}{2}+j\right)\left(\frac{k_{21}}{2}+j-1\right)}. \tag{107}$$

Combining (104) - (107), we see that

$$\begin{aligned}
\frac{\partial m_{\hat{b}_{2SLS}}(\mu'\mu; k_{21})}{\partial(\mu'\mu)} &= \left(\frac{1}{2}\right)\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\Bigg[\\
&\Bigg\{\left(\frac{1}{2}\right)\Bigg[\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{\left(\frac{k_{21}}{2}-j-2\right)}{\left(\frac{k_{21}}{2}+j\right)\left(\frac{k_{21}}{2}+j-1\right)\left(\frac{k_{21}}{2}+j-2\right)}\\
&+\sum_{j=0}^{\infty} j\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{(k_{21}-2)}{\left(\frac{k_{21}}{2}+j\right)\left(\frac{k_{21}}{2}+j-1\right)\left(\frac{k_{21}}{2}+j-2\right)}\\
&-\left(\frac{1-\rho^2}{2\rho^2}\right)\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{1}{\left(\frac{k_{21}}{2}+j\right)\left(\frac{k_{21}}{2}+j-1\right)}\\
&-2\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{\left(\frac{k_{21}}{2}-1\right)}{\left(\frac{k_{21}}{2}+j\right)\left(\frac{k_{21}}{2}+j-1\right)}\Bigg\}\\
&= \left(\frac{1}{2}\right)\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\Bigg[\\
&\Bigg\{\left(\frac{1}{2}\right)\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{\left(\frac{k_{21}}{2}+j-2\right)}{\left(\frac{k_{21}}{2}+j\right)\left(\frac{k_{21}}{2}+j-1\right)\left(\frac{k_{21}}{2}+j-2\right)}\\
&+\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{(k_{21}-2)j}{\left(\frac{k_{21}}{2}+j\right)\left(\frac{k_{21}}{2}+j-1\right)\left(\frac{k_{21}}{2}+j-2\right)}\\
&-\left(\frac{1-\rho^2}{2\rho^2}\right)\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{1}{\left(\frac{k_{21}}{2}+j\right)\left(\frac{k_{21}}{2}+j-1\right)}\\
&-2\sum_{j=0}^{\infty}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{\left(\frac{k_{21}}{2}-1\right)\left(\frac{k_{21}}{2}+j-2\right)}{\left(\frac{k_{21}}{2}+j\right)\left(\frac{k_{21}}{2}+j-1\right)\left(\frac{k_{21}}{2}+j-2\right)}\Bigg\}\\
&= \left(\frac{1}{2}\right)\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\Bigg[\\
&\Bigg\{\left(\frac{1}{2}\right)\frac{1}{\left(\frac{k_{21}}{2}\right)\left(\frac{k_{21}}{2}-1\right)}+
\end{aligned}$$

$$
\begin{aligned}
&\left(\frac{1}{2}\right)\sum_{j=1}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{(\frac{k_{21}}{2}-j-2)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}\\
&+\sum_{j=0}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{(k_{21}-2)j}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}\\
&-\left(\frac{1-\rho^2}{2\rho^2}\right)\sum_{j=0}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{1}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)}\\
&-2\sum_{j=0}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{(\frac{k_{21}}{2}-1)j}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}\\
&-2\frac{1}{(\frac{k_{21}}{2})}-2\sum_{j=1}\frac{\left(\frac{\mu'\mu}{2}\right)^j(\frac{k_{21}}{2}-1)(\frac{k_{21}}{2}-2)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}\Bigg\}\\
=\;&\left(\frac{1}{2}\right)\sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\times\\
&\Bigg\{-\frac{(k_{21}-\frac{5}{2})}{(\frac{k_{21}}{2})(\frac{k_{21}}{2}-1)}\\
&-\left(\frac{1}{2}\right)\sum_{j=1}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{(k_{21}-\frac{13}{2}k_{21}+j+10)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)(\frac{k_{21}}{2}+j-2)}\\
&-\left(\frac{1-\rho^2}{2\rho^2}\right)\sum_{j=0}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\frac{1}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)}\Bigg\}. \qquad (108)
\end{aligned}
$$

Note that $(1-\rho^2)/2\rho^2 > 0$ for $\sigma_{uv} \neq 0$ and for integer $k_{21} \geq 4$ and for all positive integer j, $(k_{21}^2 - \frac{13}{2}k_{21} + j + 10) > 0$ so all three terms inside the curly brackets of expression (108) are negative. Hence, we deduce that

$$
\frac{\partial m_{\hat{b}_{IV}}(\mu'\mu; k_{21})}{\partial(\mu'\mu)} < 0, \qquad (109)
$$

which proves part (d). $\Box$

Proof of Theorem 4.1: To show part (a), note that direct application of part (a) of Lemma A7 to the bias expression (36) yield

$$
\begin{aligned}
b_{\hat{b}_{IV}}(\tau^2; k_{21}) \;=\;& \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\\
&\times\Bigg\{\left(\frac{1}{1+\tau^2}\right) - \left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\Bigg[2-4\left(\frac{1}{1+\tau^2}\right)
\end{aligned}
$$

$$+ 2\left(\frac{1}{1+\tau^2}\right)^2\Bigg] + \left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^2\Bigg[8 + 12\left(\frac{1}{1+\tau^2}\right) - 32\left(\frac{1}{1+\tau^2}\right)^2 + 12\left(\frac{1}{1+\tau^2}\right)^3\Bigg]\Bigg\} + O(k_{21}^{-3})$$

$$= \sigma_{uu}^{1=2}\sigma_{vv}^{-1=2}\rho\left\{\left(\frac{1}{1+\tau^2}\right) - \left(\frac{2}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2\right\} + O(k_{21}^{-2}); \qquad (110)$$

which completes our proof of part (a).

To show part (b), we first rewrite expression (37) as

$$\begin{aligned} m_{\hat{b}_{IV}}(\tau^2;k_{21}) = {} & \sigma_{uu}\sigma_{vv}^{-1}\rho^2 \\ & \times\Bigg[\frac{1}{\rho^2}\left(\frac{1}{k_{21}-2}\right){}_1F_1\left(\frac{k_{21}}{2}-1;\frac{k_{21}}{2};\frac{\tau^2k_{21}}{2}\right)e^{-\frac{\tau^2k_{21}}{2}} \\ & + \left(\frac{k_{21}-3}{k_{21}-2}\right)\left(\frac{k_{21}-2}{2}\right){}_1F_1\left(\frac{k_{21}}{2}-2;\frac{k_{21}}{2}-1;\frac{\tau^2k_{21}}{2}\right)e^{-\frac{\tau^2k_{21}}{2}} \\ & - \left(\frac{k_{21}-3}{k_{21}-2}\right)\left(\frac{k_{21}-4}{2}\right){}_1F_1\left(\frac{k_{21}}{2}-1;\frac{k_{21}}{2};\frac{\tau^2k_{21}}{2}\right)e^{-\frac{\tau^2k_{21}}{2}}\Bigg]; \end{aligned} \qquad (111)$$

where we have made use of the identity

$$\begin{aligned} (\gamma-\alpha-1)\,{}_1F_1(\alpha;\gamma;z) = {} & (\gamma-1)\,{}_1F_1(\alpha;\gamma-1;z) \\ & -\alpha\,{}_1F_1(\alpha+1;\gamma;z) \end{aligned} \qquad (112)$$

in rewriting expression (37). (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 (111) above, we obtain

$$\begin{aligned} m_{\hat{b}_{IV}}(\tau^2;k_{21}) = {} & \sigma_{uu}\sigma_{vv}^{-1}\rho^2\Bigg[\frac{1}{\rho^2}\frac{1}{k_{21}}\left(\frac{1}{1-\frac{2}{k_{21}}}\right)\rho\Bigg\{\left(\frac{1}{1+\tau^2}\right) - 2\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right) \\ & + 4\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)^2 - 2\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)^3 + 8\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^2 \\ & + 12\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^3 - 32\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^4 + \end{aligned}$$

$$
\begin{aligned}
&12\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^5\Bigg\} + \left(\frac{1-\frac{3}{k_{21}}}{1-\frac{2}{k_{21}}}\right)\left(\frac{k_{21}}{2}-1\right)\Bigg\{\left(\frac{1}{1+\tau^2}\right)\\
&-4\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)+6\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)^2-2\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)^3\\
&+12\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^2+20\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^3-\\
&44\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^4+12\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^5\Bigg\}-\left(\frac{1-\frac{3}{k_{21}}}{1-\frac{2}{k_{21}}}\right)\\
&\times\left(\frac{k_{21}}{2}-2\right)\Bigg\{\left(\frac{1}{1+\tau^2}\right)-2\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)+4\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)^2\\
&-2\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)^3+8\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^2+\\
&12\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^3-32\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^4\\
&+12\left(\frac{1}{k_{21}}\right)^2\left(\frac{1}{1+\tau^2}\right)^5\Bigg\}\Bigg]+O(k_{21}^{-3})
\end{aligned}
\tag{113}
$$

After some tedious and straightforward calculations, it is possible to show that

$$
\begin{aligned}
m_{\hat{b}_{IV}}(\tau^2;k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\Bigg[\left(\frac{1}{1+\tau^2}\right)^2+\frac{1}{\rho^2}\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\\
&\quad 3\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)^2+2\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)^3\\
&\quad -6\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)^4\Bigg]+O(k_{21}^{-2})\\
&= \sigma_{uu}\sigma_{vv}^{-1}\rho^2\Bigg[\left(\frac{1}{1+\tau^2}\right)^2+\left(\frac{1-\rho^2}{\rho^2}\right)\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\\
&\quad +\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\left(1+3\left(\frac{1}{1+\tau^2}\right)+2\left(\frac{1}{1+\tau^2}\right)^2\right.\\
&\quad \left.-6\left(\frac{1}{1+\tau^2}\right)^3\right)\Bigg]+O(k_{21}^{-2});
\end{aligned}
\tag{114}
$$

which proves part (b). □

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

$$
\hat{m}_{\hat{b}_{IV}}(\tau^2;k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left(\frac{1}{1+\tau^2}\right)^2
$$

$$
\begin{aligned}
&+\sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right) \\
&+\sigma_{uu}\sigma_{vv}^{-1}\rho^2\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\cdot\left[1+3\left(\frac{1}{1+\tau^2}\right)\right. \\
&\left.+2\left(\frac{1}{1+\tau^2}\right)^2-6\left(\frac{1}{1+\tau^2}\right)^3\right].
\end{aligned}
\tag{115}
$$

Note that to show $\hat{m}_{\hat{b}_{IV}}(\tau^2;k_{21}) \geq 0$, it suffices to show that

$$
\left[1+3\left(\frac{1}{1+\tau^2}\right)+2\left(\frac{1}{1+\tau^2}\right)^2-6\left(\frac{1}{1+\tau^2}\right)^3\right] \geq 0 \tag{116}
$$

for $\tau^2 \in [0,\infty)$. To proceed, let

$$
x=\left(\frac{1}{1+\tau^2}\right) \tag{117}
$$

and define

$$
\varphi(x)=1+3x+2x^2-6x^3. \tag{118}
$$

Note further that to show the inequality condition (116), it is sufficient to show that $\varphi(x) \geq 0$ for $x \in [0,1]$. Observe first that $\varphi(0)=1 \geq 0$ and $\varphi(1)=0$, so that $\varphi(\cdot)$ is non-negative at the end points of the interval. Next, we need to analyze how $\varphi(x)$ behaves in the open interval $(0,1)$. To do so, we take the derivative of φ with respect to x to obtain

$$
\varphi'(x)=3+4x-18x^2. \tag{119}
$$

Note that $\varphi'(x)$ is a continuous function which is 0 at $x=1/9\left(1+\sqrt{58}/2\right) \approx .534$, positive for $x \in \left[0, 1/9\left(1+\sqrt{58}/2\right)\right)$, and negative for $x \in \left(1/9\left(1+\sqrt{58}/2\right), 1\right]$. It follows that $\varphi(x)$ is monotonically increasing for $x \in \left[0, 1/9\left(1+\sqrt{58}/2\right)\right)$ reaching a local maximum at $x=1/9\left(1+\sqrt{58}/2\right)$. Moreover, $\varphi(x)$ is monotonically decreasing for $x \in \left(1/9\left(1+\sqrt{58}/2\right), 1\right]$ reaching a value of 0 at $x=1$. Part (a), thus, follows immediately.

To prove (b), write

$$
\begin{aligned}
\hat{V}_{\hat{b}_{IV}}(\tau^2;k_{21}) &= \sigma_{uu}\sigma_{vv}^{-1}(1-\rho^2)\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right) \\
&+\sigma_{uu}\sigma_{vv}^{-1}\rho^2\frac{1}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\cdot\left(1+7\left(\frac{1}{1+\tau^2}\right)\right.
\end{aligned}
$$

$$-6\left(\frac{1}{1+\tau^2}\right)^2 - 2\left(\frac{1}{1+\tau^2}\right)^3\Bigg]. \tag{120}$$

Note that to show $V_{\hat{b}_{IV}}(\tau^2; k_{21}) \geq 0$, it suffices to show that

$$\left[1 + 7\left(\frac{1}{1+\tau^2}\right) - 6\left(\frac{1}{1+\tau^2}\right)^2 - 2\left(\frac{1}{1+\tau^2}\right)^3\right] \geq 0 \tag{121}$$

for $\tau^2 \in [0, \infty)$. Again, to proceed, let

$$x = \left(\frac{1}{1+\tau^2}\right) \tag{122}$$

and define

$$\theta(x) = 1 + 7x - 6x^2 - 2x^3, \tag{123}$$

so that to show the inequality condition (121), it is sufficient to show that $\theta(x) \geq 0$ for $x \in [0, 1]$. Observe first that $\theta(0) = 1 \geq 0$ and $\theta(1) = 0$, so that $\theta(\cdot)$ is non-negative at the end points of the interval. Next, we need to analyze how $\theta(x)$ behaves in the open interval $(0, 1)$,. To do so, we take the derivative of θ with respect to x to obtain

$$\theta'(x) = 7 - 12x - 6x^2. \tag{124}$$

Note that $\theta'(x)$ is a continuous function which is 0 at $x = \sqrt{78}/6 - 1 \approx .472$, positive for $x \in \left[0, \sqrt{78}/6 - 1\right)$, and negative for $x \in \left(\sqrt{78}/6 - 1, 1\right]$. It follows that $\theta(x)$ is monotonically increasing for $x \in \left[0, \sqrt{78}/6 - 1\right)$ reaching a local maximum at $x = \sqrt{78}/6 - 1$. Moreover, $\theta(x)$ is monotonically decreasing for $x \in \left(\sqrt{78}/6 - 1, 1\right]$ reaching a value of 0 at $x = 1$. Hence, part (b) is also proved. □

Proof of Theorem 4.4: To show part (a), we note that we can rewrite the bias expression (36) in the alternative form

$$\begin{aligned} b_{\hat{b}_{IV}}(\tau^2; k_{21}) &= \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \rho \left[1 - e^{-\frac{\tau^2 k_{21}}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\tau^2 k_{21}}{2}\right)^{j+1}}{j!} \frac{1}{\left(\frac{k_{21}}{2} + j\right)}\right] \\ &= \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} \rho f(\tau^2; k_{21}) \quad \text{(say)}. \end{aligned} \tag{125}$$

Taking the partial derivative of $f(\tau^2; k_{21})$ in (125) above with respect to k_{21}, we obtain

$$
\begin{aligned}
\frac{\partial f(\tau^2;k_{21})}{\partial k_{21}} = e^{-\frac{\tau^2 k_{21}}{2}} \Bigg[& \left(\frac{\tau^2}{2}\right) \sum_{j=0}^{\infty} \frac{\left(\frac{\tau^2 k_{21}}{2}\right)^{j+1}}{j!} \frac{1}{(\frac{k_{21}}{2}+j)} \\
& + \left(\frac{1}{2}\right) \sum_{j=0}^{\infty} \frac{\left(\frac{\tau^2 k_{21}}{2}\right)^{j+1}}{j!} \frac{1}{(\frac{k_{21}}{2}+j)^2} \\
& - \left(\frac{\tau^2}{2}\right) \sum_{j=0}^{\infty} \frac{(j+1)\left(\frac{\tau^2 k_{21}}{2}\right)^{j}}{j!} \frac{1}{(\frac{k_{21}}{2}+j)};
\end{aligned} \tag{126}
$$

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

$$
\begin{aligned}
\frac{\partial f(\tau^2;k_{21})}{\partial k_{21}} &= e^{-\frac{\tau^2 k_{21}}{2}} \Bigg[\left(\frac{\tau^2}{2}\right) \sum_{i=0}^{\infty} \frac{i\left(\frac{\tau^2 k_{21}}{2}\right)^{i}}{i!} \frac{(\frac{k_{21}}{2}+i)}{(\frac{k_{21}}{2}+i)(\frac{k_{21}}{2}+i-1)} \\
&\quad - \left(\frac{\tau^2}{2}\right) \sum_{j=0}^{\infty} \frac{j\left(\frac{\tau^2 k_{21}}{2}\right)^{j}}{j!} \frac{(\frac{k_{21}}{2}+j-1)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)} \\
&\quad - \left(\frac{\tau^2}{2}\right) \sum_{j=0}^{\infty} \frac{\left(\frac{\tau^2 k_{21}}{2}\right)^{j}}{j!} \frac{(\frac{k_{21}}{2}+j-1)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)} \\
&\quad + \left(\frac{1}{2}\right) \sum_{j=0}^{\infty} \frac{\left(\frac{\tau^2 k_{21}}{2}\right)^{j+1}}{j!} \frac{1}{(\frac{k_{21}}{2}+j)^2} \Bigg] \\
&= e^{-\frac{\tau^2 k_{21}}{2}} \Bigg[-\left(\frac{\tau^2}{2}\right) \sum_{j=0}^{\infty} \frac{\left(\frac{\tau^2 k_{21}}{2}\right)^{j}}{j!} \frac{(\frac{k_{21}}{2}-1)}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)} \\
&\quad + \left(\frac{1}{2}\right) \sum_{j=0}^{\infty} \frac{\left(\frac{\tau^2 k_{21}}{2}\right)^{j+1}}{j!} \frac{1}{(\frac{k_{21}}{2}+j)^2} \Bigg] \\
&= e^{-\frac{\tau^2 k_{21}}{2}} \Bigg[\sum_{j=0}^{\infty} \frac{\left(\frac{\tau^2 k_{21}}{2}\right)^{j}}{j!} \frac{\gamma_j}{(\frac{k_{21}}{2}+j)(\frac{k_{21}}{2}+j-1)} \Bigg] > 0
\end{aligned} \tag{127}
$$

Since $f(\tau^2;k_{21})$ is nonnegative, as shown in the proof of part (d) of Theorem 3.1, it follows that

$$
\left| b_{IV}(\tau^2;k_{21}) \right| = |\rho| \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} f(\tau^2;k_{21}) \tag{128}
$$

and, thus, $\left|\hat{\beta}_{\hat{\mathbf{b}}_{IV}}(\tau^2; k_{21})\right|$ is a monotonically increasing function of k_{21} for τ^2 fixed. Part (a) is, thus, proved.

Finally, we note that part (b) is true if

$$\left[1+3\left(\frac{1}{1+\tau^2}\right)+2\left(\frac{1}{1+\tau^2}\right)^2-6\left(\frac{1}{1+\tau^2}\right)^3\right] \geq 0; \qquad (129)$$

but the inequality condition (129) was already shown to be true in the proof of part(a) of Theorem 4.3. Hence, part (b) follows immediately. $\square$

References

Angrist, J. D. and A.D. 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.

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

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

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

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. (1980). \The Exact Distribution of Instrumental Variable Estimators in an Equation Containing n+1 Endogenous Variables," Econometrica, 48, 861-878.

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. (1989). \Partially Identi¯ed Econometric Models," Econometric Theory, 5, 181-240.

Richardson, D. H. (1968). \The Exact Distribution of a Structural Coe±cient 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.

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

Slater, L.J. (1960). Con°uent 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 Coe±cients in Linear Regression," Journal of Econometrics, 2, 143-150.

Zivot, E. and J. Wang (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 $\rightarrow$ Regressor	IV Bias			IV Bias/OLS Bias		
intercept	0.003 (66.42)	-0.056 (-112.2)	-1.2E-7 (-1.2E-5)	0.002 (211.9)	-0.022 (-246.8)	0.599 (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 (76860)	1.006 (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 (53651)	1.028 (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: All regressions are run using bias figures (as dependent variable) generated according to the analytical formulae given above, and/or using least squares bias (which is equal to $(\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho$). All parameters are defined above. In addition to least squares regression parameter estimates (non-bracketed entries in the table), t-statistics are reported (bracketed entries).

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

Dependent Variable −− > Regressor	IV MSE			IV MSE/OLS MSE		
intercept	-0.128	0.029	-0.299	-0.007	0.004	0.396
	(-65.86)	(16.33)	(-3.628)	(-24.32)	(19.71)	(345.8)
$((\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho)^2$			0.396			
			(238.9)			
$((\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho)^2(\delta'\delta/k_{21})$			-0.038			
			(-104.7)			
$((\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho)^{-2}/k_{21}$						0.298
						(17.98)
$\delta'\delta/k_{21}$						-0.038
						(-126.1)
$1/k_{21}$			8.059			
			(7.543)			
$(1+\delta'\delta/k_{21})^{-2}$				1.002	1.001	
				(1472)	(1889)	
$((\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho)^2(1+\delta'\delta/k_{21})^{-2}$	0.998	0.999				
	(10350)	(10952)				
$(1+\delta'\delta/k_{21})^{-1}/k_{21}$	6.098					
	(54.89)					
$((\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho)^{-2}(1+\delta'\delta/k_{21})^{-1}/k_{21}$				-0.016		
				(-81.68)		
$(1+\delta'\delta/k_{21})^{-1}/k_{21}$				2.341		
				(47.20)		
$(1+\delta'\delta/k_{21})^{-2}/k_{21}$				-4.464		
				(-13.79)		
$(1+\delta'\delta/k_{21})^{-3}/k_{21}$				6.536		
				(10.02)		
$(1+\delta'\delta/k_{21})^{-4}/k_{21}$				-3.119		
				(-8.224)		
$((\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho)^2(1+\delta'\delta/k_{21})^{-1}/k_{21}$	1.217					
	(132.8)					
$((\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho)^2(1+\delta'\delta/k_{21})^{-2}/k_{21}$	-6.126					
	(-97.87)					
$((\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho)^2(1+\delta'\delta/k_{21})^{-3}/k_{21}$	10.51					
	(83.79)					
$((\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2})\rho)^2(1+\delta'\delta/k_{21})^{-4}/k_{21}$	-5.622					
	(-77.19)					
$\bar{R}^2$	0.9998	0.9996	0.5494	0.9893	0.9859	0.2392

[1] Notes: See notes to Table 1. All regressions are run using mse figures (as dependent variable) generated according to the analytical formulae given above, and/or using least squares mse (which is equal to $(\sigma_{uu}\sigma_{vv}^{-1})\rho^2$).