Biometrika (2015), **xx**, x, *pp.* 1–19

Printed in Great Britain

Supplementary Material for "On Distributions of Ratios"

BY SIMON A. BRODA

Department of Economics and Econometrics, University of Amsterdam, Valckenierstraat 65–67 1018XE Amsterdam, The Netherlands

s.a.broda@uva.nl

AND RAYMOND KAN

Joseph L. Rotman School of Management, University of Toronto, 105 St. George Street Toronto, Ontario, Canada M5S 3E6

kan@chass.utoronto.ca

1. PROOF THAT (8) REDUCES TO (6) IF X AND Y FORM A DEFINITE PAIR

To see that

$$\frac{1}{\pi^2}\int_0^\infty\int_{-\infty}^\infty \mathrm{Re}\left\{\varphi_2(s,-t-rs)\right\}\mathrm{d}s\,\frac{\mathrm{d}t}{t} = \left|\frac{1}{\pi}\int_0^\infty \mathrm{Im}\left\{\varphi_2(s,-rs)\right\}\mathrm{d}s\right|$$

when X and Y form a definite pair, assume that $r \neq \beta$ and rewrite the left hand side as

$$-\frac{1}{\pi^2}\,\mathrm{Re}\int_0^\infty \oint_0 \varphi_2(s,t-rs)\frac{\mathrm{d}t}{t}\,\mathrm{d}s.$$

We consider the case with $\delta = 1$ and $\beta < \infty$; the other cases can be treated analogously. Make the change of variables $s \mapsto s+t$, $t \mapsto (r-\beta)t$, so that the inner integrand becomes $f_s(t) \equiv \mathrm{sgn}(r-\beta)\varphi_2(s+t,-sr-\beta t)$. Observe that $\varphi_{X,Y}(s+t,-sr-\beta t) = E\left\{\exp(\mathrm{i}sW_r + \mathrm{i}tW_\beta)\right\}$, where $W_\beta < 0$ almost surely. This implies that for real s and as a function of t, $\varphi_{X,Y}(s+t,-sr-\beta t)$, and hence $f_s(t)$, is analytic for $\mathrm{Im}\, t < 0$. Consider a contour that consists of a line segment from $-T$ to $-1/T$, a small counterclockwise loop half way around the origin, another line segment from $1/T$ to T, and a large semicircle in the lower half of the complex plane back to $-T$. The contour encloses no singularities, hence the integral along it is zero. As $T \to \infty$, the integral along the large semicircle converges to zero. The integral along the half loop around the origin is equal to minus one half the residue at the origin, and hence

$$\oint_0 f_s(t)\frac{\mathrm{d}t}{t} = \mathrm{i}\pi f_s(0).$$

2. EXACTNESS OF THE DENSITY APPROXIMATION IN THE GAUSSIAN CASE

Suppose that X and Y are jointly Gaussian with respective means μ_X and μ_Y, variances σ_X^2 and σ_Y^2, and correlation ρ. The density of $R = X/Y$ has been found in Fieller (1932); see also Hinkley (1969). The cumulant generating function of X and Y is $\mathbb{K}(s,t) = s\mu_X + t\mu_Y + (s^2\sigma_X^2 + 2st\rho\sigma_X\sigma_Y + t^2\sigma_Y^2)/2$. Applying Theorem 3 with $n = 1$, it is found that both saddle-points are explicit in terms of the parameters. Defining

$$a \equiv \left(\frac{r^2}{\sigma_X^2} - \frac{2r\rho}{\sigma_X \sigma_Y} + \frac{1}{\sigma_Y^2} \right)^{1/2}, \quad b \equiv \frac{r\mu_X}{\sigma_X^2} - \frac{\rho(\mu_X + r\mu_Y)}{\sigma_X \sigma_Y} + \frac{\mu_Y}{\sigma_Y^2},$$

they are given by

$$\tilde{s}_0 = \frac{r\mu_Y - \mu_X}{a^2 \sigma_X^2 \sigma_Y^2}, \quad \hat{s} = \frac{\mu_Y \rho \sigma_X - \mu_X \sigma_Y}{\sigma_X^2 \sigma_Y (1-\rho^2)}, \quad \hat{t} = \frac{\mu_X \rho \sigma_Y - \mu_Y \sigma_X}{\sigma_Y^2 \sigma_X (1-\rho^2)}.$$

The other relevant quantities are

$$\tilde{g}_0 = \frac{b}{a^3 \sigma_X \sigma_Y}, \quad \tilde{w}_0 = \frac{r\mu_Y - \mu_X}{a \sigma_X \sigma_Y}, \quad \hat{w} = -\frac{b}{(1-\rho^2)^{1/2} a}.$$

Plugging in and rearranging, this is exactly the expression given in Fieller (1932) and Hinkley (1969); in other words, the saddlepoint approximation is exact in this case.

3. Non-negativity of the Density Approximation

We first show that $\tilde{g}_0$ and $\hat{w}$ in (17) have opposite signs. As a function of r, both $\hat{w}$ and $\mathbb{K}_2(\tilde{s}_0, -r\tilde{s}_0)$ switch sign only at $r = r_0$. It is immediate that $\hat{w}$ crosses the abscissa from below if $\hat{s} > 0$ and from above otherwise. Regarding $\mathbb{K}_2(\tilde{s}_0, -r\tilde{s}_0)$, differentiate (16) with respect to r to obtain

$$\tilde{s}_0' \equiv \frac{\mathrm{d}}{\mathrm{d}r}\tilde{s}_0 = \frac{\tilde{s}_0 \left\{ \mathbb{K}_{12}(\tilde{s}_0, -r\tilde{s}_0) - r\mathbb{K}_{22}(\tilde{s}_0, -r\tilde{s}_0) \right\} + \mathbb{K}_2(\tilde{s}_0, -r\tilde{s}_0)}{c_r^{\mathrm{T}} \mathbb{K}''(\tilde{s}_0, -r\tilde{s}_0) c_r}.$$

Using that $\lim_{r \to r_0} \tilde{s}_0' = \hat{s}\{\hat{\mathbb{K}}_{12}(\hat{s}, \hat{t}) - r_0 \hat{\mathbb{K}}_{22}(\hat{s}, \hat{t})\} / \{c_{r_0}^{\mathrm{T}} \hat{\mathbb{K}}''(\hat{s}, \hat{t}) c_{r_0}\}$ and simplifying,

$$\left. \frac{\mathrm{d}}{\mathrm{d}r} \mathbb{K}_2(\tilde{s}_0, -r\tilde{s}_0) \right|_{r=r_0} = -\hat{s} \frac{|\mathbb{K}''(\hat{s}, \hat{t})|}{c_{r_0}^{\mathrm{T}} \hat{\mathbb{K}}''(\hat{s}, \hat{t}) c_{r_0}},$$

so that $\mathbb{K}_2(\tilde{s}_0, -r\tilde{s}_0)$ crosses the axis in the opposite direction as $\hat{w}$ and consequently has the opposite sign. The non-negativity of (17) is then seen as follows. Consider the function $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto \Phi(x) + \phi(x)/x$. Then $f(x) < 0, x < 0$ and $f(x) > 1, x > 0$. This follows directly from Gordon (1941, Eq. 7), who shows that for $x > 0$, $\Phi(-x)/\phi(-x) < 1/x$. Thus the term in square brackets in (17) is greater than one if $\hat{w} < 0$ and smaller than minus one if $\hat{w} > 0$. The result follows because $\tilde{g}_0$ and $\hat{w}$ have opposite signs.

4. Higher Order Terms for the Density Approximation

Let $\tilde{s} = \tilde{s}(t)$ be the inner saddlepoint, that is, the solution to the equation

$$\mathbb{K}_1(\tilde{s}, t - r\tilde{s}) = r\mathbb{K}_2(\tilde{s}, t - r\tilde{s}).$$

Applying a standard Laplace approximation to $I_1(t)$ in (10) yields

$$\frac{n}{2\pi\mathrm{i}} \int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} e^{n\mathbb{K}(s, t-rs)} \mathbb{K}_2(s, t - rs) \mathrm{d}s = \frac{\sqrt{n}}{(2\pi)^{1/2}} e^{nh(t)} \left\{ \sum_{j=0}^{m-1} \frac{g_j(t)}{n^j} + \mathcal{O}(n^{-m}) \right\}, \tag{S1}$$

where $h(t) \equiv \mathbb{K}(\tilde{s}, t - r\tilde{s})$. An explicit expression for $g_j(t)$ can be obtained using Eq. (103) of Rice (1968). It is

$$g_j(t) = \sum_{k=0}^{2j} \frac{\tilde{J}_k(t)}{k!\tilde{h}_2(t)^{\frac{k+1}{2}}} \tilde{a}_{j,2j-k}(t), \tag{S2}$$

where

$$\tilde{h}_k(t) \equiv \left. \frac{\partial^k \mathbb{K}(s, t - rs)}{\partial s^k} \right|_{s=\tilde{s}} = \sum_{j=0}^{k} \binom{k}{j} (-r)^j \mathbb{K}_{1^{k-j}2^j}(\tilde{s}, t - r\tilde{s}),$$

$$\tilde{J}_k(t) \equiv \left. \frac{\partial^k \mathbb{K}_2(s, t - rs)}{\partial s^k} \right|_{s=\tilde{s}} = \sum_{j=0}^{k} \binom{k}{j} (-r)^j \mathbb{K}_{1^{k-j}2^{j+1}}(\tilde{s}, t - r\tilde{s}),$$

$\mathbb{K}_{1^i 2^j}(s,t) \equiv \partial^{i+j}\mathbb{K}(s,t)/(\partial s^i \partial t^j)$, the coefficients $\tilde{a}_{i,j}(t)$ are given by

$$\tilde{a}_{i,j}(t) \equiv \sum_{k=0}^{j} \tilde{d}_{k,j}(t)(-2)^{i+k} \left(\frac{1}{2}\right)_{i+k},$$

$(x)_k$ denotes the rising factorial, the $\tilde{d}_{i,j}(t)$ are obtained from the recurrence relation $\tilde{d}_{0,0}(t) = 1$, $\tilde{d}_{0,j}(t) = 0$ $(j \geq 1)$, and

$$\tilde{d}_{i,j}(t) = \frac{1}{i} \sum_{k=1}^{j-i+1} \frac{\tilde{h}_{k+2}(t)}{(k+2)!\tilde{h}_2(t)^{\frac{k+2}{2}}} \tilde{d}_{i-1,j-k}(t) \quad (j \geq i \geq 1).$$

Using (S1), one has

$$I_2 = \frac{\sqrt{n}}{(2\pi)^{1/2}} \frac{1}{2\pi \mathrm{i}} \int_{c_2 - \mathrm{i}\infty}^{c_2 + \mathrm{i}\infty} e^{nh(t)} \left\{ \sum_{j=0}^{m-1} \frac{g_j(t)}{n^j} + \mathcal{O}(n^{-m}) \right\} \frac{\mathrm{d}t}{t}.$$

Denote a typical term in the integral as

$$I_{2,j} \equiv \frac{\sqrt{n}}{(2\pi)^{1/2}} \frac{1}{2\pi \mathrm{i}} \int_{c_2 - \mathrm{i}\infty}^{c_2 + \mathrm{i}\infty} e^{nh(t)} g_j(t) \frac{\mathrm{d}t}{t}.$$

Using the result from Bleistein (1966) and Rice (1968, Appendix F, by setting $\lambda = 0$), $I_{2,j}$ can be approximated as

$$I_{2,j} = \frac{\sqrt{n}}{(2\pi)^{1/2}} e^{nh(0)} \left[\{1_{c_2>0} - \Phi(\hat{w}\sqrt{n})\} g_j(0) + \frac{\phi(\hat{w}\sqrt{n})}{\sqrt{n}} \left\{ \sum_{k=0}^{m-1} \frac{p_{j,k}}{n^k} + \mathcal{O}(n^{-m}) \right\} \right],$$

where $\hat{t}_r \equiv \hat{t} + r\hat{s}$, $\hat{w} \equiv \mathrm{sgn}(\hat{t}_r) \left[2\left\{h(0) - h(\hat{t}_r)\right\}\right]^{1/2}$, $\hat{u} \equiv \hat{t}_r\{h''(\hat{t}_r)\}^{1/2}$,

$$p_{j,k} = \sum_{l=0}^{2k} \sum_{q=0}^{l} \frac{(-1)^l g_j^{(q)}(\hat{t}_r)}{q! h''(\hat{t}_r)^{\frac{q}{2}} \hat{u}^{2k+1-l}} a_{k,l-q} - \frac{g_j(0)(-2)^k \left(\frac{1}{2}\right)_k}{\hat{w}^{2k+1}}, \tag{S3}$$

and the coefficients $a_{i,j}$ are given by

$$a_{i,j} \equiv \sum_{k=0}^{j} d_{k,j}(-2)^{i+k} \left(\frac{1}{2}\right)_{i+k},$$

with $d_{i,j}$ obtained from the recurrence relation

$$d_{0,0} = 1, \qquad d_{0,j} = 0,\ j \geq 1, \qquad d_{i,j} = \frac{1}{i} \sum_{k=1}^{j-i+1} \theta_{k+2} d_{i-1,j-k} \quad (j \geq i \geq 1).$$

Here $\theta_k \equiv h^{(k)}(\hat{t}_r)/\{k! h''(\hat{t}_r)^{k/2}\}$. Details of the derivation of this formula are available upon request. Collecting the terms with like power of n, the mth order approximation for I_2 is

$$I_2 = \sqrt{n}\phi(\tilde{w}_0\sqrt{n}) \left[\{1_{c_2>0} - \Phi(\hat{w}\sqrt{n})\} \sum_{j=0}^{m-1} \frac{g_j(0)}{n^j} + \frac{\phi(\hat{w}\sqrt{n})}{\sqrt{n}} \sum_{j=0}^{m-1} \frac{1}{n^j} \sum_{k=0}^{j} p_{j-k,k} + \mathcal{O}(n^{-m}) \right].$$

Similarly, the mth order approximation for $I_1(0)$ is

$$I_1(0) = \frac{n}{2\pi \mathrm{i}} \int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} e^{n\mathbb{K}(s,-rs)} \mathbb{K}_2(s,-rs) \mathrm{d}s = \sqrt{n}\phi(\tilde{w}_0\sqrt{n}) \left[\sum_{j=0}^{m-1} \frac{g_j(0)}{n^j} + \mathcal{O}(n^{-m}) \right].$$

It follows that the mth order saddlepoint approximation for the density is

$$\hat{f}_n^{(m)}(r) = \sqrt{n}\phi(\tilde{w}_0\sqrt{n}) \left[\{1 - 2\Phi(\hat{w}\sqrt{n})\} \sum_{j=0}^{m-1} \frac{g_j(0)}{n^j} + \frac{2\phi(\hat{w}\sqrt{n})}{\sqrt{n}} \sum_{j=0}^{m-1} \frac{A_j}{n^j} \right],$$

where

$$\begin{aligned}
A_j &\equiv \sum_{k=0}^{j} p_{j-k,k} \\
&= \sum_{k=0}^{j} \sum_{l=0}^{2k} \sum_{q=0}^{l} \frac{(-1)^l g_{j-k}^{(q)}(\hat{t}_r)}{q! h''(\hat{t}_r)^{\frac{q}{2}} \hat{u}^{2k+1-l}} a_{k,l-q} - \sum_{k=0}^{j} \frac{g_{j-k}(0)(-2)^k \left(\frac{1}{2}\right)_k}{\hat{w}^{2k+1}} \\
&= \sum_{k=0}^{j} \sum_{l=0}^{2k} \frac{(-1)^l}{\hat{u}^{2k+1-l}} b_{j,k,l} - \sum_{k=0}^{j} \frac{g_{j-k}(0)(-2)^k \left(\frac{1}{2}\right)_k}{\hat{w}^{2k+1}},
\end{aligned}$$

and

$$b_{j,k,l} \equiv \sum_{q=0}^{l} \frac{g_{j-k}^{(q)}(\hat{t}_r) a_{k,l-q}}{q! h''(\hat{t}_r)^{\frac{q}{2}}}.$$

The above expression for A_j is undefined when $\hat{t}_r = 0$, i.e., when $\hat{t} = -r\hat{s}$. In order to obtain its limit as $\hat{t}_r \to 0$, expand h in a Taylor series about $\hat{t}_r$. This yields

$$h(0) - h(\hat{t}_r) = -\hat{t}_r h'(\hat{t}_r) + \frac{\hat{t}_r^2}{2} h''(\hat{t}_r) - \frac{\hat{t}_r^3}{3!} h'''(\hat{t}_r) + \frac{\hat{t}_r^4}{4!} h^{(4)}(\hat{t}_r) - \cdots.$$

Using the fact that $h'(\hat{t}_r) = 0$, one has

$$\begin{aligned}\hat{w}^2 &= 2\left\{\mathbb{K}(\tilde{s}_0, -r\tilde{s}_0) - \mathbb{K}(\hat{s}, \hat{t})\right\}\\ &= 2\left\{h(0) - h(\hat{t}_r)\right\}\\ &= 2\left\{\frac{\hat{t}_r^2}{2!}h''(\hat{t}_r) - \frac{\hat{t}_r^3}{3!}h'''(\hat{t}_r) + \frac{\hat{t}_r^4}{4!}h^{(4)}(\hat{t}_r) - \cdots\right\}\\ &= 2\sum_{j=2}^{\infty}(-1)^j\theta_j\hat{u}^j.\end{aligned}$$

Letting

$$B_{j,k} \equiv \frac{g_j(0)(-2)^k\left(\frac{1}{2}\right)_k}{\hat{w}^{2k+1}}$$

and using (S3),

$$\begin{aligned}B_{j,k} &= \sum_{l=0}^{2k}\sum_{q=0}^{l}\frac{(-1)^l g_j^{(q)}(\hat{t}_r)}{q!h''(\hat{t}_r)^{\frac{q}{2}}\hat{u}^{2k+1-l}}a_{k,l-q} - p_{j,k},\\ B_{j,k+1} &= \sum_{l=0}^{2k+2}\sum_{q=0}^{l}\frac{(-1)^l g_j^{(q)}(\hat{t}_r)}{q!h''(\hat{t}_r)^{\frac{q}{2}}\hat{u}^{2k+3-l}}a_{k+1,l-q} - p_{j,k+1}.\end{aligned}$$

Using the fact that $\hat{w}^2 B_{j,k+1} = -(2k+1)B_{j,k}$, it follows that

$$\begin{aligned}&\left\{2\sum_{j=2}^{\infty}(-1)^j\theta_j\hat{u}^j\right\}\left\{\sum_{l=0}^{2k+2}\sum_{q=0}^{l}\frac{(-1)^l g_j^{(q)}(\hat{t}_r)}{q!h''(\hat{t}_r)^{\frac{q}{2}}\hat{u}^{2k+3-l}}a_{k+1,l-q} - p_{j,k+1}\right\}\\ &\qquad = -(2k+1)\left\{\sum_{l=0}^{2k}\sum_{q=0}^{l}\frac{(-1)^r g_j^{(q)}(\hat{t}_r)}{q!h''(\hat{t}_r)^{\frac{q}{2}}\hat{u}^{2k+1-l}}a_{k,l-q} - p_{j,k}\right\}.\end{aligned}$$

Comparing the constant term on both sides, $p_{j,k}$ can be expressed as

$$p_{j,k} = -\frac{2}{2k+1}\sum_{l=0}^{2k+1}\theta_{2k+3-l}\sum_{q=0}^{l}\frac{g_j^{(q)}(\hat{t}_r)a_{k+1,l-q}}{q!h''(\hat{t}_r)^{\frac{q}{2}}}.$$

Taking the limit,

$$\lim_{\hat{t}_r\to 0} p_{j,k} = -\frac{2}{2k+1}\sum_{l=0}^{2k+1}\bar{\theta}_{2k+3-l}\sum_{q=0}^{l}\frac{g_j^{(q)}(0)\bar{a}_{k+1,l-q}}{q!h''(0)^{\frac{q}{2}}},$$

where $\bar{\theta}_j$ and $\bar{a}_{i,j}$ are the values of θ_j and $a_{i,j}$ evaluated at $\hat{t}_r = 0$. It follows that at $r = r_0 \equiv -\hat{t}/\hat{s}$ where $\hat{t}_r = 0$,

$$\hat{f}_n^{(m)}(r_0) = (2/\pi)^{1/2}\phi(\tilde{w}_0\sqrt{n})\sum_{j=0}^{m-1}\frac{\bar{A}_j}{n^j},$$

where

$$\begin{aligned}\bar{A}_j &\equiv \lim_{\hat{t}_r \to 0} A_j \\ &= \lim_{\hat{t}_r \to 0} \sum_{k=0}^{j} p_{j-k,k} \\ &= -\sum_{k=0}^{j} \frac{2}{2k+1} \sum_{l=0}^{2k+1} \bar{\theta}_{2k+3-l} \sum_{q=0}^{l} \frac{g_{j-k}^{(q)}(0)\bar{a}_{k+1,l-q}}{q! h''(0)^{\frac{q}{2}}} \\ &= -\sum_{k=0}^{j} \frac{2}{2k+1} \sum_{l=0}^{2k+1} \bar{\theta}_{2k+3-l} \bar{b}_{j+1,k+1,l},\end{aligned}$$

and $\bar{b}_{j+1,k+1,l}$ is the value of $b_{j+1,k+1,l}$ evaluated at $\hat{t}_r = 0$. The following subsections provide explicit expressions obtained by specializing these results to the first, second, and third order cases.

4·1. *First Order Approximation*

For $m = 1$ and $\hat{t}_r \neq 0$, using that $g_0(\hat{t}_r) = 0$ yields

$$A_0 = p_{0,0} = \frac{g_0(\hat{t}_r)}{\hat{u}} - \frac{g_0(0)}{\hat{w}} = -\frac{g_0(0)}{\hat{w}}.$$

It follows that when $r \neq r_0$ so that $\hat{t}_r \neq 0$,

$$\hat{f}_n^{(1)}(r) = \sqrt{n}\phi(\tilde{w}_0\sqrt{n}) g_0(0) \left[1 - 2\Phi(\hat{w}\sqrt{n}) - \frac{2\phi(\hat{w}\sqrt{n})}{\hat{w}\sqrt{n}}\right],$$

which yields (17). Regarding the limit at $\hat{t}_r = 0$,

$$\bar{A}_0 = -\bar{\theta}_3 g_0(0) + \frac{g_0'(0)}{h''(0)^{\frac{1}{2}}} = \frac{g_0'(0)}{h''(0)^{\frac{1}{2}}} = \frac{h''(0)^{\frac{1}{2}}}{\tilde{h}_2(0)^{\frac{1}{2}}} = \frac{|\mathbb{K}''(\hat{s},\hat{t})|^{\frac{1}{2}}}{c_{r_0}^{\mathrm{T}} \mathbb{K}''(\hat{s},\hat{t}) c_{r_0}}.$$

It follows that at $r = r_0$,

$$\hat{f}_n^{(1)}(r_0) = (2/\pi)^{1/2} \phi(\tilde{w}_0\sqrt{n}) \frac{|\mathbb{K}''(\hat{s},\hat{t})|^{\frac{1}{2}}}{c_{r_0}^{\mathrm{T}} \mathbb{K}''(\hat{s},\hat{t}) c_{r_0}},$$

which proves (18).

4·2. *Second Order Approximation*

For $m = 2$ and $\hat{t}_r \neq 0$, one has

$$\hat{f}_n^{(2)}(r) = \hat{f}_n^{(1)}(r) + \frac{\phi(\tilde{w}_0\sqrt{n}) g_1(0)}{\sqrt{n}} \{1 - 2\Phi(\hat{w}\sqrt{n})\} + \frac{2\phi(\tilde{w}_0\sqrt{n})\phi(\hat{w}\sqrt{n})}{n} A_1,$$

where

$$\begin{aligned}A_1 &= p_{0,1} + p_{1,0} \\ &= \frac{g_1(\hat{t}_r)}{\hat{u}} - \frac{g_0''(\hat{t}_r)}{2h''(\hat{t}_r)\hat{u}} + \frac{3\theta_3 g_0'(\hat{t}_r)}{h''(\hat{t}_r)^{\frac{1}{2}}\hat{u}} + \frac{g_0'(\hat{t}_r)}{h''(\hat{t}_r)^{\frac{1}{2}}\hat{u}^2} + \frac{g_0(0)}{\hat{w}^3} - \frac{g_1(0)}{\hat{w}}.\end{aligned}$$

Writing $g_0(t) \equiv h'(t)/\tilde{h}_2(t)^{1/2}$, it is easy to see that

$$g_0'(\hat{t}_r) = \frac{h''(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)^{\frac{1}{2}}}, \quad g_0''(\hat{t}_r) = \frac{h'''(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)^{\frac{1}{2}}} - \frac{h''(\hat{t}_r)\tilde{h}_2'(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}} = \frac{6\theta_3 h''(\hat{t}_r)^{\frac{3}{2}}}{\tilde{h}_2(\hat{t}_r)^{\frac{1}{2}}} - \frac{h''(\hat{t}_r)\tilde{h}_2'(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}}.$$

Using these expressions, one has that

$$\frac{g_0''(\hat{t}_r)}{2h''(\hat{t}_r)} - \frac{3\theta_3 g_0'(\hat{t}_r)}{h''(\hat{t}_r)^{\frac{1}{2}}} = -\frac{\tilde{h}_2'(\hat{t}_r)}{2\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}}. \tag{S4}$$

Using (S2) and the fact that $\tilde{J}_0(t) = h'(t)$, $g_1(t)$ can be written as

$$\begin{aligned} g_1(t) &= \left\{ \frac{\tilde{h}_4(t)}{8\tilde{h}_2(t)^2} - \frac{5\tilde{h}_3(t)^2}{24\tilde{h}_2(t)^3} \right\} \frac{h'(t)}{\tilde{h}_2(t)^{\frac{1}{2}}} + \frac{1}{2\tilde{h}_2(t)^{\frac{3}{2}}} \left\{ \frac{\tilde{h}_3(t)\tilde{J}_1(t)}{\tilde{h}_2(t)} - \tilde{J}_2(t) \right\} \\ &= \left\{ \frac{\tilde{h}_4(t)}{8\tilde{h}_2(t)^2} - \frac{5\tilde{h}_3(t)^2}{24\tilde{h}_2(t)^3} \right\} \frac{h'(t)}{\tilde{h}_2(t)^{\frac{1}{2}}} - \frac{\tilde{h}_2'(t)}{2\tilde{h}_2(t)^{\frac{3}{2}}}, \end{aligned} \tag{S5}$$

where the last equality follows from the identity

$$\tilde{h}_k'(t) = \tilde{h}_{k+1}(t)\tilde{s}'(t) + \tilde{J}_k(t) = -\frac{\tilde{h}_{k+1}(t)\tilde{J}_1(t)}{\tilde{h}_2(t)} + \tilde{J}_k(t) \quad (k \geq 1). \tag{S6}$$

As $h'(\hat{t}_r) = 0$, $g_1(\hat{t}_r)$ reduces to

$$g_1(\hat{t}_r) = -\frac{\tilde{h}_2'(\hat{t}_r)}{2\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}}. \tag{S7}$$

Using (S4) and (S7), A_1 can be simplified to

$$A_1 = \frac{g_0'(\hat{t}_r)}{h''(\hat{t}_r)^{\frac{1}{2}}\hat{u}^2} + \frac{g_0(0)}{\hat{w}^3} - \frac{g_1(0)}{\hat{w}} = \frac{1}{\hat{t}_r^2 |\mathbb{K}''(\hat{s}, \hat{t})|^{\frac{1}{2}}} + \frac{g_0(0)}{\hat{w}^3} - \frac{g_1(0)}{\hat{w}}.$$

Regarding the limit at $\hat{t}_r = 0$,

$$\hat{f}_n^{(2)}(r_0) = (2/\pi)^{1/2} \phi(\tilde{w}_0 \sqrt{n}) \left\{ \frac{|\mathbb{K}(\hat{s}, \hat{t})|^{\frac{1}{2}}}{c_{r_0}^{\mathrm{T}} \mathbb{K}''(\hat{s}, \hat{t}) c_{r_0}} + \frac{\bar{A}_1}{n} \right\},$$

where

$$\bar{A}_1 = \left(3\bar{\theta}_4 - \frac{15\bar{\theta}_3^2}{2} \right) \frac{g_0'(0)}{h''(0)^{\frac{1}{2}}} + \frac{3\bar{\theta}_3 g_0''(0)}{2h''(0)} - \frac{g_0'''(0)}{6h''(0)^{\frac{3}{2}}} - \bar{\theta}_3 g_1(0) + \frac{g_1'(0)}{h''(0)^{\frac{1}{2}}}.$$

Evaluating $\bar{A}_1$ requires explicit expressions for $g_0'''(\hat{t}_r)$, $g_1'(\hat{t}_r)$, $h'''(\hat{t}_r)$ and $h^{(4)}(\hat{t}_r)$. It is straightforward to show that

$$\begin{aligned} g_0'''(t) = \frac{h^{(4)}(t)}{\tilde{h}_2(t)^{\frac{1}{2}}} - \frac{3h'''(t)\tilde{h}_2'(t)}{2\tilde{h}_2(t)^{\frac{3}{2}}} + \frac{9h''(t)\tilde{h}_2'(t)^2}{4\tilde{h}_2(t)^{\frac{5}{2}}} - \frac{3h''(t)\tilde{h}_2''(t)}{2\tilde{h}_2(t)^{\frac{3}{2}}} \\ + h'(t) \left\{ -\frac{15\tilde{h}_2'(t)^3}{8\tilde{h}_2(t)^{\frac{7}{2}}} + \frac{9\tilde{h}_2'(t)\tilde{h}_2''(t)}{4\tilde{h}_2(t)^{\frac{5}{2}}} - \frac{\tilde{h}_2'''(t)}{2\tilde{h}_2(t)^{\frac{3}{2}}} \right\}. \end{aligned}$$

Thus, again using the fact that $h'(\hat{t}_r) = 0$,

$$\begin{aligned} g_0'''(\hat{t}_r) &= \frac{h^{(4)}(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)^{\frac{1}{2}}} - \frac{3h'''(\hat{t}_r)\tilde{h}_2'(\hat{t}_r)}{2\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}} + \frac{9h''(\hat{t}_r)\tilde{h}_2'(\hat{t}_r)^2}{4\tilde{h}_2(\hat{t}_r)^{\frac{5}{2}}} - \frac{3h''(\hat{t}_r)\tilde{h}_2''(\hat{t}_r)}{2\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}} \\ &= \frac{24\theta_4 h''(\hat{t}_r)^2}{\tilde{h}_2(\hat{t}_r)^{\frac{1}{2}}} - \frac{9\theta_3 h''(\hat{t}_r)^{\frac{3}{2}}\tilde{h}_2'(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}} + \frac{9h''(\hat{t}_r)\tilde{h}_2'(\hat{t}_r)^2}{4\tilde{h}_2(\hat{t}_r)^{\frac{5}{2}}} - \frac{3h''(\hat{t}_r)\tilde{h}_2''(\hat{t}_r)}{2\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}}, \end{aligned}$$

where $\tilde{h}_2'(\hat{t}_r)$ and $\tilde{h}_2''(\hat{t}_r)$ are obtained from (S6) and given by

$$\begin{aligned} \tilde{h}_2'(\hat{t}_r) &= \tilde{J}_2(\hat{t}_r) - \frac{\tilde{h}_3(\hat{t}_r)\tilde{J}_1(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)}, \\ \tilde{h}_2''(\hat{t}_r) &= \tilde{J}_2'(\hat{t}_r) - \frac{\tilde{h}_3'(\hat{t}_r)\tilde{J}_1(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)} - \frac{\tilde{h}_3(\hat{t}_r)\tilde{J}_1'(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)} + \frac{\tilde{h}_3(\hat{t}_r)\tilde{J}_1(\hat{t}_r)\tilde{h}_2'(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)^2}, \end{aligned}$$

$\tilde{h}_3'(\hat{t}_r)$ can be obtained from (S6), and the general expression for $\tilde{J}_k'(t)$ is

$$\tilde{J}_k'(t) = -\frac{\tilde{J}_{k+1}(t)\tilde{J}_1(t)}{\tilde{h}_2(t)} + \sum_{j=0}^{k} \binom{k}{j} (-r)^j \mathbb{K}_{1^{k-j}2^{j+2}}(\tilde{s}, t - r\tilde{s}).$$

Differentiating (S5) and using the fact that $h'(\hat{t}_r) = 0$ once more, it is found that

$$g_1'(\hat{t}_r) = \left\{ \frac{\tilde{h}_4(\hat{t}_r)}{8\tilde{h}_2(\hat{t}_r)^2} - \frac{5\tilde{h}_3(\hat{t}_r)^2}{24\tilde{h}_2(\hat{t}_r)^3} \right\} \frac{h''(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)^{\frac{1}{2}}} - \frac{\tilde{h}_2''(\hat{t}_r)}{2\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}} + \frac{3\tilde{h}_2'(\hat{t}_r)^2}{4\tilde{h}_2(\hat{t}_r)^{\frac{5}{2}}}.$$

Regarding $h'''(t)$ and $h^{(4)}(t)$, using that $h''(t) = \mathbb{K}_{22}(\tilde{s}, t - r\tilde{s}) - \{\tilde{J}_1(t)^2/\tilde{h}_2(t)\}$ yields

$$\begin{aligned} h'''(t) &= \mathbb{K}_{222}(\tilde{s}, t - r\tilde{s}) + \{\mathbb{K}_{122}(\tilde{s}, t - r\tilde{s}) - r\mathbb{K}_{222}(\tilde{s}, t - r\tilde{s})\}\, \tilde{s}'(t) \\ &\quad - \frac{2\tilde{J}_1(t)\tilde{J}_1'(t)}{\tilde{h}_2(t)} + \frac{\tilde{J}_1(t)^2\tilde{h}_2'(t)}{\tilde{h}_2(t)^2} \\ &= \mathbb{K}_{222}(\tilde{s}, t - r\tilde{s}) - 3\{\mathbb{K}_{122}(\tilde{s}, t - r\tilde{s}) - r\mathbb{K}_{222}(\tilde{s}, t - r\tilde{s})\} \frac{\tilde{J}_1(t)}{\tilde{h}_2(t)} \\ &\quad + \frac{3\tilde{J}_1(t)^2\tilde{J}_2(t)}{\tilde{h}_2(t)^2} - \frac{\tilde{J}_1(t)^3\tilde{h}_3(t)}{\tilde{h}_2(t)^3}, \\ h^{(4)}(t) &= \mathbb{K}_{2^4}(\tilde{s}, t - r\tilde{s}) - 4\{\mathbb{K}_{12^3}(\tilde{s}, t - r\tilde{s}) - r\mathbb{K}_{2^4}(\tilde{s}, t - r\tilde{s})\} \frac{\tilde{J}_1(t)}{\tilde{h}_2(t)} \\ &\quad + 6\left\{\mathbb{K}_{1^22^2}(\tilde{s}, t - r\tilde{s}) - 2r\mathbb{K}_{12^3}(\tilde{s}, t - r\tilde{s}) + r^2\mathbb{K}_{2^4}(\tilde{s}, t - r\tilde{s})\right\} \frac{\tilde{J}_1(t)^2}{\tilde{h}_2(t)^2} \\ &\quad - \frac{3\left\{\tilde{J}_1'(t)\tilde{h}_2(t) - \tilde{J}_1(t)\tilde{h}_2'(t)\right\}^2}{\tilde{h}_2(t)^3} - \frac{4\tilde{J}_1(t)^3\tilde{J}_3(t)}{\tilde{h}_2(t)^3} + \frac{\tilde{J}_1(t)^4\tilde{h}_4(t)}{\tilde{h}_2(t)^4}. \end{aligned}$$

Setting $\hat{t}_r = 0$ yields

$$\bar{A}_1 = \left\{ \frac{3\bar{\theta}_3^2 - 2\bar{\theta}_4}{2} + \frac{\tilde{h}_4(0)}{8\tilde{h}_2(0)^2} - \frac{5\tilde{h}_3(0)^2}{24\tilde{h}_2(0)^3} \right\} \frac{|\mathbb{K}(\hat{s},\hat{t})|^{\frac{1}{2}}}{c_{r_0}^{\mathrm{T}}\mathbb{K}''(\hat{s},\hat{t})c_{r_0}} + \frac{\bar{\theta}_3 \tilde{h}_2'(0)}{2\tilde{h}_2(0)^{\frac{3}{2}}} - \frac{\tilde{h}_2''(0)}{4\tilde{h}_2(0)^{\frac{3}{2}} h''(0)^{\frac{1}{2}}} + \frac{3\tilde{h}_2'(0)^2}{8\tilde{h}_2(0)^{\frac{5}{2}} h''(0)^{\frac{1}{2}}}.$$

4·3. *Third Order Approximation*

For $m = 3$ and $\hat{t}_r \neq 0$, one has

$$\hat{f}_n^{(3)}(r) = \hat{f}_n^{(2)}(r) + \frac{\phi(\tilde{w}_0\sqrt{n})g_2(0)}{n^{\frac{3}{2}}}\{1 - 2\Phi(\hat{w}\sqrt{n})\} + \frac{2\phi(\tilde{w}_0\sqrt{n})\phi(\hat{w}\sqrt{n})}{n^2} A_2,$$

where

$$\begin{aligned} A_2 &= p_{2,0} + p_{1,1} + p_{0,2} \\ &= \frac{g_2(\hat{t}_r)}{\hat{u}} - \frac{g_2(0)}{\hat{w}} + \left(\frac{a_{1,0}}{\hat{u}^3} - \frac{a_{1,1}}{\hat{u}^2} + \frac{a_{1,2}}{\hat{u}}\right) g_1(\hat{t}_r) \\ &\quad - \left(\frac{a_{1,0}}{\hat{u}^2} - \frac{a_{1,1}}{\hat{u}}\right) \frac{g_1'(\hat{t}_r)}{h''(\hat{t}_r)^{\frac{1}{2}}} + \frac{a_{1,0}}{\hat{u}} \frac{g_1''(\hat{t}_r)}{2h''(\hat{t}_r)} + \frac{g_1(0)}{\hat{w}^3} \\ &\quad - \left(\frac{a_{2,0}}{\hat{u}^4} - \frac{a_{2,1}}{\hat{u}^3} + \frac{a_{2,2}}{\hat{u}^2} - \frac{a_{2,3}}{\hat{u}}\right) \frac{g_0'(\hat{t}_r)}{h''(\hat{t}_r)^{\frac{1}{2}}} \\ &\quad + \left(\frac{a_{2,0}}{\hat{u}^3} - \frac{a_{2,1}}{\hat{u}^2} + \frac{a_{2,2}}{\hat{u}}\right) \frac{g_0''(\hat{t}_r)}{2h''(\hat{t}_r)} - \left(\frac{a_{2,0}}{\hat{u}^2} - \frac{a_{2,1}}{\hat{u}}\right) \frac{g_0'''(\hat{t}_r)}{6h''(\hat{t}_r)^{\frac{3}{2}}} \\ &\quad + \frac{a_{2,0}}{\hat{u}} \frac{g_0^{(4)}(\hat{t}_r)}{24h''(\hat{t}_r)^2} - \frac{3g_0(0)}{\hat{w}^5}. \end{aligned}$$

Using (S6) and the identity

$$\sum_{k=1}^{2j} \frac{\tilde{h}_{k+1}(t)}{k!\tilde{h}_2(t)^{\frac{k+1}{2}}} \tilde{a}_{j,2j-k}(t) = 0,$$

$g_j(\hat{t}_r)$ can be written as

$$g_j(\hat{t}_r) = \sum_{k=2}^{2j} \frac{\tilde{h}_k'(\hat{t}_r)}{k!\tilde{h}_2(\hat{t}_r)^{\frac{k+1}{2}}} \tilde{a}_{j,2j-k}(\hat{t}_r).$$

Specifically,

$$\begin{aligned} g_2(\hat{t}_r) &= \frac{\tilde{a}_{2,2}(\hat{t}_r)\tilde{h}_2'(\hat{t}_r)}{2\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}} + \frac{\tilde{a}_{2,1}(\hat{t}_r)\tilde{h}_3'(\hat{t}_r)}{6\tilde{h}_2(\hat{t}_r)^2} + \frac{\tilde{a}_{2,0}(\hat{t}_r)\tilde{h}_4'(\hat{t}_r)}{24\tilde{h}_2(\hat{t}_r)^{\frac{5}{2}}} \\ &= \frac{35\tilde{h}_3(\hat{t}_r)^2\tilde{h}_2'(\hat{t}_r) - 20\tilde{h}_2(\hat{t}_r)h_3(\hat{t}_r)\tilde{h}_3'(\hat{t}_r)}{48\tilde{h}_2(\hat{t}_r)^{\frac{9}{2}}} - \frac{15\tilde{h}_2(\hat{t}_r)\tilde{h}_4(\hat{t}_r)\tilde{h}_2'(\hat{t}_r) + 6\tilde{h}_2(\hat{t}_r)^2\tilde{h}_4'(\hat{t}_r)}{48\tilde{h}_2(\hat{t}_r)^{\frac{9}{2}}}. \end{aligned}$$

After some simplification, it can be verified that the coefficient associated with $1/\hat{u}$ in A_2 is zero. Therefore, A_2 can be written as

$$
\begin{aligned}
A_2 = \frac{1}{\hat{u}^2}\Bigg[&\left\{3\theta_4 - \frac{15\theta_3^2}{2} + \frac{\tilde{h}_4(\hat{t}_r)}{8\tilde{h}_2(\hat{t}_r)^2} - \frac{5\tilde{h}_3(\hat{t}_r)^2}{24\tilde{h}_2(\hat{t}_r)^2}\right\}\frac{h''(\hat{t}_r)^{\frac{1}{2}}}{\tilde{h}_2(\hat{t}_r)^{\frac{1}{2}}} \\
&- \frac{3\theta_3\tilde{h}_2'(\hat{t}_r)}{2\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}} + \frac{\tilde{h}_2''(\hat{t}_r)}{4\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}h''(\hat{t}_r)^{\frac{1}{2}}} - \frac{3\tilde{h}_2'(\hat{t}_r)^2}{8\tilde{h}_2(\hat{t}_r)^{\frac{5}{2}}h''(\hat{t}_r)^{\frac{1}{2}}}\Bigg] \\
&- \frac{1}{\hat{u}^3}\left\{\frac{\tilde{h}_2'(\hat{t}_r)}{\tilde{h}_2(\hat{t}_r)^{\frac{3}{2}}} + \frac{6\theta_3 h''(\hat{t}_r)^{\frac{1}{2}}}{\tilde{h}_2(\hat{t}_r)^{\frac{1}{2}}}\right\} - \frac{3\tilde{h}_2''(\hat{t}_r)^{\frac{1}{2}}}{\hat{u}^4\tilde{h}_2(\hat{t}_r)^{\frac{1}{2}}} \\
&- \frac{g_2(0)}{\hat{w}} + \frac{g_1(0)}{\hat{w}^3} - \frac{3g_0(0)}{\hat{w}^5}.
\end{aligned}
$$

Regarding the limit at $\hat{t}_r = 0$,

$$
\hat{f}_n^{(3)}(r_0) = (2/\pi)^{1/2}\phi(\tilde{w}_0\sqrt{n})\left\{\frac{|\mathbb{K}(\hat{s},\hat{t})|^{\frac{1}{2}}}{c_{r_0}^{\mathrm{T}}\mathbb{K}''(\hat{s},\hat{t})c_{r_0}} + \frac{\bar{A}_1}{n} + \frac{\bar{A}_2}{n^2}\right\},
$$

where

$$
\begin{aligned}
\bar{A}_2 = &\frac{g_0^{(5)}(0)}{40h''(0)^{\frac{5}{2}}} - \frac{5\bar{\theta}_3 g_0^{(4)}(0)}{8h''(0)^2} + \frac{5(7\bar{\theta}_3^2 - 2\bar{\theta}_4)g_0'''(0)}{4h''(0)^{\frac{3}{2}}} \\
&- \frac{15(21\bar{\theta}_3^3 - 14\bar{\theta}_3\bar{\theta}_4 + 2\bar{\theta}_5)g_0''(0)}{4h''(0)} \\
&+ \frac{15(231\bar{\theta}_3^4 - 252\bar{\theta}_3^2\bar{\theta}_4 + 28\bar{\theta}_4^2 + 56\bar{\theta}_3\bar{\theta}_5 - 8\bar{\theta}_6)g_0'(0)}{8h''(0)^{\frac{1}{2}}} \\
&+ \left(3\bar{\theta}_4 - \frac{15\bar{\theta}_3^2}{2}\right)\frac{g_1'(0)}{h''(0)^{\frac{1}{2}}} + \frac{3\bar{\theta}_3 g_1''(0)}{2h''(0)} - \frac{g_1'''(0)}{6h''(0)^{\frac{3}{2}}} - \bar{\theta}_3 g_2(0) + \frac{g_2'(0)}{h''(0)^{\frac{1}{2}}}.
\end{aligned}
$$

Evaluating $\bar{A}_2$ requires explicit expressions for $h^{(5)}(t)$ and $h^{(6)}(t)$. These can be obtained by differentiating $h^{(4)}(t)$, but they are lengthy and hence omitted here.

5. Failure of the Method of Proof used for the Density when Applied to Tail Probabilities

We briefly discuss why the method of proof used in the derivation of the density approximation fails for the distribution function. Consider Equation (11). For the two univariate probabilities, obtaining a uniform asymptotic expansion is a simple task. All that is required is an application of Lemma 4 to the relevant inversion integrals. Doing so leads to the tail probability approximation of Lugannani & Rice (1980). Approximating the bivariate probability is less straightforward. The relevant inversion formula reads, for $c_1, c_2 < 0$,

$$
F_{\bar{W},\bar{Y}}^n(0,0) = \left(\frac{1}{2\pi\mathrm{i}}\right)^2 \int_{c_2-\mathrm{i}\infty}^{c_2+\mathrm{i}\infty}\int_{c_1-\mathrm{i}\infty}^{c_1+\mathrm{i}\infty} e^{n\mathbb{K}(s,t-rs)}\frac{\mathrm{d}s}{s}\frac{\mathrm{d}t}{t}. \tag{S8}
$$

In order to appreciate the difficulties involved in approximating (S8), it is useful to compare it to the double integral I_2 in (10) whose uniform asymptotic expansion formed the basis for the density approximation (17). The essential difference is the presence of the pole in the inner integrand. Applying a standard Laplace approximation as in (13) would therefore result in an expansion which is nonuniform in r as the saddlepoint crosses the pole. Instead, the inner integral could be approximated by another application of Lemma 4; this is the approach taken by Wang (1990) in deriving a saddlepoint approximation for bivariate distributions. Unfortunately, when applied to the present problem, the approximation contains a term $\mathbb{K}(0, \hat{t} + r\hat{s})$, or w_{u_0} in Wang's notation. Depending on $\mathbb{K}(\cdot, \cdot)$, $(0, \hat{t} + r\hat{s})$ may fall outside the convergence region $\mathcal{T}$ for some values of r, rendering the approximation invalid. Although not discussed by Wang, this problem occurs not only in the present context, but more generally in the approximation of a bivariate distribution function, the subject of his paper. Kolassa & Li (2010) develop an alternative to Wang's approximation which is also applicable in higher dimensional problems, but it suffers from the same deficiency (see also Li, 2009, in particular Eq. 3.2.3).

6. Joint Moment Generating Function for the Application in Section 5·1

6·1. *Derivation*

Let $Z_1, Z_2 \sim \mathrm{N}(0, 1)$ and $X_i \sim \chi^2_{\nu_i}, i \in \{1, 2\}$, with respective density

$$f_{\nu_i}(x) = \frac{e^{-x/2} x^{(\nu_i - 2)/2}}{2^{\nu_i/2}\Gamma(\nu_i/2)},$$

all independent. The joint moment generating function of

$$X \equiv [a\{X_1X_2/(\nu_1\nu_2)\}^{1/2} + bZ_1(X_2/\nu_2)^{1/2}], Y \equiv [c\{X_1X_2/(\nu_1\nu_2)\}^{1/2} + dZ_2(X_1/\nu_1)^{1/2}]$$

is, after integrating Z_1, Z_2,

$$\mathbb{M}(s, t) = \int_0^\infty \int_0^\infty \exp\left\{\frac{d^2t^2x_1}{2\nu_1} + \frac{(as + ct)(x_1x_2)^{1/2}}{(\nu_1\nu_2)^{1/2}} + \frac{b^2s^2x_2}{2\nu_2}\right\} f_{\nu_2}(x_2) f_{\nu_1}(x_1)\mathrm{d}x_2\mathrm{d}x_1,$$

defined for all (s, t) such that

$$-\sqrt{\nu_2}/|b| < s < \sqrt{\nu_2}/|b|, \qquad \max\left[-\sqrt{\nu_1}/|d|, M^-\right] < t < \min\left[\sqrt{\nu_1}/|d|, M^+\right],$$

where

$$M^\pm \equiv \frac{-acs \pm \{(c^2 + d^2\omega_2)\nu_1\omega_2 - a^2d^2s^2\omega_2\}^{1/2}}{c^2 + d^2\omega_2}.$$

From Gradshteyn & Ryzhik (2007, 3.462.1), one has that for $\alpha < 1$ and with $y = -\beta(1 - \alpha)^{-1/2}$,

$$\int_0^\infty e^{\alpha u/2 + \beta\sqrt{u}} f_\nu(u)\mathrm{d}u = \frac{\Gamma(\nu)e^{y^2/4}}{\Gamma(\nu/2)(1-\alpha)^{\nu/2}2^{(\nu-2)/2}} \mathrm{D}_{-\nu}(y),$$

where

$$\mathrm{D}_{-\nu}(y) \equiv \frac{e^{-y^2/4}}{\Gamma(\nu)} \int_0^\infty x^{\nu-1} e^{-x^2/2 - xy}\mathrm{d}x$$

is the parabolic cylinder function. Hence, with $\omega_2 \equiv= \nu_2 - b^2s^2$ and $y \equiv -(as + ct)\{x_1/(\nu_1\omega_2)\}^{1/2}$,

$$\mathbb{M}(s, t) = \left(\frac{\nu_2}{\omega_2}\right)^{\nu_2/2} \frac{\Gamma(\nu_2)}{2^{(\nu_2-2)/2}\Gamma(\nu_2/2)} \int_0^\infty \exp\left(\frac{d^2t^2x_1}{2\nu_1}\right) e^{y^2/4}\mathrm{D}_{-\nu_2}(y) f_{\nu_1}(x_1)\mathrm{d}x_1.$$

Using the identity

$$\mathrm{D}_{-\nu}(y) = 2^{-\nu/2} e^{-y^2/4} \sqrt{\pi} \left[\frac{1}{\Gamma\{(1+\nu)/2\}} {}_1\mathrm{F}_1\left(\frac{\nu}{2}; \frac{1}{2}; \frac{y^2}{2}\right) - \frac{y\sqrt{2}}{\Gamma(\nu/2)} {}_1\mathrm{F}_1\left(\frac{1+\nu}{2}; \frac{3}{2}; \frac{y^2}{2}\right) \right]$$

(Gradshteyn & Ryzhik, 2007, 9·240), where ${}_1\mathrm{F}_1$ is the confluent hypergeometric function, one obtains

$$\begin{aligned} \mathbb{M}(s,t) = {} & \frac{(\nu_2/\omega_2)^{\nu_2/2}}{2^{\nu_1/2}\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \int_0^\infty x_1^{(\nu_1-2)/2} e^{-\omega_1 x_1/(2\nu_1)} \\ & \times \left\{ \Gamma\left(\frac{\nu_2}{2}\right) {}_1\mathrm{F}_1\left(\frac{\nu_2}{2}; \frac{1}{2}; \frac{y^2}{2}\right) - y\sqrt{2}\Gamma\left(\frac{\nu_2+1}{2}\right) {}_1\mathrm{F}_1\left(\frac{\nu_2+1}{2}; \frac{3}{2}; \frac{y^2}{2}\right) \right\} \mathrm{d}x_1, \end{aligned}$$

where $\omega_1 \equiv \nu_1 - d^2 t^2$. Next, for $|s| > |k|$, $\mathrm{Re}\,\beta > 0$, $\mathrm{Re}\,s > \max(0, \mathrm{Re}\,k)$, one has

$$\int_0^\infty e^{-st} t^{\beta-1} {}_1\mathrm{F}_1(\alpha; \gamma; kt) = \Gamma(b) s^{-b} {}_2\mathrm{F}_1(\alpha, \beta; \gamma; k/s),$$

(Gradshteyn & Ryzhik, 2007, 7.621.4), so that

$$\begin{aligned} \mathbb{M}(s,t) = {} & \left(\frac{\nu_1}{\omega_1}\right)^{\nu_1/2} \left(\frac{\nu_2}{\omega_2}\right)^{\nu_2/2} \left[{}_2\mathrm{F}_1\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}; \frac{1}{2}; z^2\right) \right. \\ & \left. + \frac{2z\Gamma\{(\nu_1+1)/2\}\Gamma\{(\nu_2+1)/2\}}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} {}_2\mathrm{F}_1\left(\frac{\nu_1+1}{2}, \frac{\nu_2+1}{2}; \frac{3}{2}; z^2\right) \right], \end{aligned}$$

where $z \equiv (as+ct)(\omega_1\omega_2)^{-1/2}$. Using Gradshteyn & Ryzhik (2007, 9.136.2), this can be simplified to

$$\begin{aligned} \mathbb{M}(s,t) = {} & \left(\frac{\nu_1}{\omega_1}\right)^{\nu_1/2} \left(\frac{\nu_2}{\omega_2}\right)^{\nu_2/2} \frac{\Gamma\{(\nu_1+1)/2\}\Gamma\{(\nu_2+1)/2\}}{\Gamma(1/2)\Gamma\{(\nu_1+\nu_2+1)/2\}} \\ & \times {}_2\mathrm{F}_1\left(\nu_1, \nu_2; \frac{\nu_1+\nu_2+1}{2}; \frac{1+z}{2}\right). \end{aligned}$$

6·2. *Evaluation*

The Gauss hypergeometric function is defined by the power series

$$ {}_2\mathrm{F}_1(a,b;c;z) \equiv \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!}, \qquad z \in \mathbb{C}, \quad |z| < 1, \quad -c \notin \mathbb{N}, \tag{S9}$$

where $(x)_n$ denotes the rising factorial. The power series converges for $|z| < 1$, but the function can be analytically continued beyond this disk. In the special case required here where $a = \nu_1$ and $b = \nu_2$ are nonnegative integers and $c = (\nu_1 + \nu_2 + 1)/2$, the function can be evaluated as a finite sum, but numerical issues arise in certain cases.

Let $\nu \equiv \min(\nu_1, \nu_2)$ and $2n \equiv \max(\nu_1, \nu_2) - \nu$. As ${}_2\mathrm{F}_1(a,b;c;z) = {}_2\mathrm{F}_1(b,a;c;z)$,

$${}_2\mathrm{F}_1\left(\nu_1, \nu_2; \frac{\nu_1+\nu_2+1}{2}; \frac{1+z}{2}\right) = {}_2\mathrm{F}_1\left(\nu+2n, \nu; \nu+n+1/2; \frac{1+z}{2}\right) \equiv f_\nu^n(z).$$

In the special case $z = 0$, one has

$$f_\nu^n(z) = \sqrt{\pi} \frac{\Gamma(\nu+n+1/2)\Gamma\{(\nu+1)/2+n\}}{\Gamma\{(\nu+1)/2\}}.$$

In general, if n is a half integer, then

$$f_\nu^n(z) = \left(\frac{2}{1-z}\right)^{\frac{\nu_1+\nu_2-1}{2}} \sum_{k=0}^{m} \frac{(-m)_k (m+1)_k \left(\frac{1+z}{2}\right)^k}{\left(\frac{\nu_1+\nu_2+1}{2}\right)_k k!} \tag{S10}$$

$$= \left(\frac{2}{1-z}\right)^{\nu_2} \sum_{k=0}^{m} \frac{(-m)_k (\nu_2)_k \left(\frac{z+1}{z-1}\right)^k}{\left(\frac{\nu_1+\nu_2+1}{2}\right)_k k!}, \tag{S11}$$

where $m \equiv n - 1/2$. These follow from the transformation formulae

$$\begin{aligned} {}_2\mathrm{F}_1(a,b;c;x) &= (1-x)^{c-a-b}{}_2\mathrm{F}_1(c-a,c-b;c;x) \\ &= (1-x)^{-b}{}_2\mathrm{F}_1\left(c-a,b;c;\frac{x}{x-1}\right). \end{aligned}$$

The first identity is numerically stable for $z \le -1$, and the second for $-1 < z < 1$. When $z \to 1$, $f_\nu^n(z)$ diverges to infinity. When $z > 1$, both (S10) and (S11) are numerically unstable, but this region is irrelevant for our application. For completeness, the following recursive expression can be used in that case. Let

$$g_k \equiv {}_2\mathrm{F}_1\left(-k, m+1; \nu_2+m+1; \frac{1+z}{2}\right),$$

so that

$$f_\nu^n(z) = \left(\frac{2}{1-z}\right)^{\nu_2+m} g_m.$$

The recurrence relations for the Gauss hypergeometric function imply that

$$g_k = \frac{1}{\nu_2+m+k}\left[\left\{m+\nu_2+2k-1-\frac{(m+k)(1+z)}{2}\right\} g_{k-1} + \frac{k(z-1)}{z+1} g_{k-2}\right],$$

with boundary conditions

$$g_0 = 1, \qquad g_1 = 1 - \frac{(m+1)(1+z)}{2(\nu_2+m+1)}.$$

If n is an integer, let

$$g_k \equiv {}_2\mathrm{F}_1\left(k+2n, k; k+n+\frac{1}{2}; \frac{1+z}{2}\right),$$

so that $f_\nu^n(z) = g_\nu$. The differential equation defining the Gauss hypergeometric function implies the recursive relationship

$$g_k = \frac{(2n+2k-1)(2n+2k-3)}{(2n+k-1)(k-1)(1-z^2)}(z g_{k-1} + g_{k-2}), \tag{S12}$$

where $g_0 = 1$ and $g_1 = h_n$, with $h_k \equiv {}_2F_1\left\{2k+1, 1; k+3/2; (1+z)/2\right\}$ given by the recursion

$$h_k = \frac{\left(k+\frac{1}{2}\right)(z+h_{k-1})}{k(1-z^2)}, \qquad h_0 = \frac{\arccos(-z)}{(1-z^2)^{1/2}}.$$

Recursion (S12) is numerically stable if $0 < z < 1$. Outside of this interval, $f_\nu^n(z)$ can be computed by summation of power series. When $-1 \le z < 0$, (S9) can be used directly, as

$0 \leq (1+z)/2 < 1/2$ implies that the series converges rapidly. In all other cases, the transformation theory of Forrey (1997) should be applied. Specifically, when $-3 \leq z < -1$, one uses the transformation

$$f_\nu^n(z) = \left(\frac{2}{1-z}\right)^{\nu_1} {}_2\mathrm{F}_1\left(\nu_1, n+\frac{1}{2}; \frac{\nu_1+\nu_2+1}{2}; w\right), \tag{S13}$$

where $w = (z+1)/(z-1)$. When $z < -3$, one uses

$$\begin{aligned} f_\nu^n(z) = {} & \frac{\Gamma\left(\nu+n+\frac{1}{2}\right)}{\Gamma(\nu_1)} \sum_{k=0}^{2n-1} \frac{\Gamma(2n-k)(\nu)_k w^{\nu+k}}{\Gamma\left(n+\frac{1}{2}-k\right)k!} \\ & + \frac{\Gamma\left(\nu+n+\frac{1}{2}\right)}{\Gamma(\nu)\Gamma\left(\frac{1}{2}-n\right)} \sum_{k=0}^{\infty} \frac{(\nu_1)_k \left(n+\frac{1}{2}\right)_k w^{\nu+2n+k}}{(2n+k)!k!} g_k, \end{aligned} \tag{S14}$$

where $w = 2/(1-z)$,

$$g_k = -\log(w) + \psi(k+1) + \psi(2n+k+1) - \psi(\nu_1+k) - \psi\left(n+k+1/2\right),$$

$\psi(k)$ is the digamma function, and g_k satisfies the recurrence relation

$$g_k = g_{k-1} + \frac{1}{k} + \frac{1}{2n+k} - \frac{1}{\nu+2n+k-1} - \frac{1}{n+\frac{1}{2}+k-1}.$$

Expression (S14) is numerically stable if $z \leq z_0 \equiv \min\{-3, (w_0-2)/w_0\}$, where $w_0 \equiv \exp\{\psi(1) + \psi(2n+1) - \psi(\nu_1) - \psi\left(n+1/2\right)\}$. When $z_0 < z < -1$, one reverts to (S13).

For $z = 1$, ones has $f_\nu^n(1) = (-1)^v$. The case with $z > 1$ is not relevant in our setting, but we discuss it for completeness. When $1 < z \leq 3$, one uses the transformation

$$\begin{aligned} & f_\nu^n(z) = (-1)^\nu (1-w)^{\nu_1} {}_2F_1\left(\nu_1, n+\frac{1}{2}; \nu+n+\frac{1}{2}; w\right) + \\ & \frac{\Gamma\left(\nu+n+\frac{1}{2}\right)\Gamma\left(\nu+n-\frac{1}{2}\right)}{\Gamma(\nu_1)\Gamma(\nu)} \mathrm{i}(-1)^{\nu+n}(1-w)^{\nu_1} w^{\frac{1}{2}-\nu-n} \sum_{k=0}^{\nu-1} \frac{(1-\nu)_k \left(n+\frac{1}{2}\right)_k w^k}{\left(-\nu-n+\frac{3}{2}\right)_k k!}, \end{aligned} \tag{S15}$$

where $w = (z-1)/(z+1)$. The first term corresponds to the real part of $f_\nu^n(z)$ and the second term to the imaginary part.

Finally when $z > 3$, one uses the transformation

$$\begin{aligned} f_\nu^n(z) = {} & \frac{(-1)^\nu \Gamma\left(\nu+n+\frac{1}{2}\right)}{\Gamma(\nu_1)} \sum_{k=0}^{2n-1} \frac{\Gamma(2n-k)(\nu)_k w^{\nu+k}}{\Gamma\left(n+\frac{1}{2}-k\right)k!} \\ & + \frac{(-1)^\nu \Gamma\left(\nu+n+\frac{1}{2}\right)}{\Gamma(\nu)\Gamma\left(\frac{1}{2}-n\right)} \sum_{k=0}^{\infty} \frac{(\nu_1)_k \left(n+\frac{1}{2}\right)_k w^{\nu+2n+k}}{(2n+k)!k!} g_k, \end{aligned}$$

where $w = 2/(1+z)$,

$$\begin{aligned} g_k &= \ln(-w^{-1}) + \psi(k+1) + \psi(2n+k+1) - \psi(\nu_1+k) - \psi\left(n+k+\frac{1}{2}\right) \\ &= \pi\mathrm{i} - \ln(w) + \psi(k+1) + \psi(2n+k+1) - \psi(\nu_1+k) - \psi\left(n+k+\frac{1}{2}\right), \end{aligned}$$

and g_k follows the same recurrence relation as in (6·2). The imaginary part $\mathrm{Im}\{f_\nu^n(z)\}$ can be simplified to

$$\frac{(-1)^{n+\nu}\Gamma\left(\nu+n+\frac{1}{2}\right)\Gamma\left(n+\frac{1}{2}\right)w^{\nu_1}(1-w)^{-n-\frac{1}{2}}}{\Gamma(\nu)\Gamma(2n+1)}\sum_{k=0}^{\nu-1}\frac{(1-\nu)_k\left(n+\frac{1}{2}\right)_k}{(2n+1)_k k!}\left(\frac{2}{1-z}\right)^k,$$

but the real part of g_k can change sign when k increases if $z<-z_0$, which leads to cancellation error, so one falls back to (S15) in that case.

7. Extension of the Application in Section 5·2: Bootstrap Inference in a Simultaneous Equations Model

Section 5·2 of the paper showed how the saddlepoint approximation can be used to construct bootstrap inferences in the Fieller–Creasy problem. A closely related question concerns inference for structural parameters in a system of simultaneous equations; see Forchini & Hillier (2003) for a discussion of the relation between the two situations. Simultaneous equation models are useful for obtaining causal inferences when randomized controlled trials cannot be conducted, as is often the case in econometrics. They have also gained some notoriety in epidemiology, specifally in the context of Mendelian randomization, as evidenced by the recent textbook of Burgess & Thompson (2015). Here, we consider a system containing just one endogenous regressor and one external instrument. In this setting, the results of the paper can be used to evaluate the distribution function of the two stage least squares estimator, on which inference can then be based. The finite sample distribution of the estimator under normality has been studied intensively; see, e.g., Richardson (1968), Sawa (1969), Anderson & Sawa (1973), Holly & Phillips (1979), Nelson & Startz (1990a,b), and Maddala & Jeong (1992), or more recently Woglom (2001), Hillier (2006), Forchini (2006), and Phillips (2006). Few authors have considered the distribution under non-Gaussianity. Knight (1986) assumes that the error distribution permits an Edgeworth expansion. Forchini (2007) considers spherically distributed errors, and Broda (2013) assumes a multivariate generalized hyperbolic distribution. The results derived herein allow us to evaluate the distribution function for a large class of error distributions. Here we follow Davidson & MacKinnon (2010) and focus on the bootstrap, extending the example in Section 5·2.

Consider the just identified model with one endogenous regressor,

$$y_1 = y_2\beta + X\gamma + u, \tag{S16}$$

$$y_2 = z_1\pi + X\delta + v, \tag{S17}$$

where $y_1\equiv(y_{1,1},\ldots,y_{1,T})^{\mathrm T}$, $y_2\equiv(y_{2,1},\ldots,y_{2,T})^{\mathrm T}$, $u\equiv(u_1,\ldots,u_T)^{\mathrm T}$, $v\equiv(v_1,\ldots,v_T)^{\mathrm T}$, X is a $T\times k$ matrix of exogenous regressors, the $T\times 1$ vector $z_1\equiv(z_{1,1},\ldots,z_{1,T})^{\mathrm T}$ represents an external instrument, $(z_1\ X)$ has full column rank, and (u_i,v_i) is independent across i and identically distributed with mean zero, $E(u_i^2)=\sigma_u^2$, $E(v_i^2)=\sigma_v^2$, and $E(u_iv_i)=\rho\sigma_u\sigma_v$, $\sigma_u,\sigma_v>0$, $|\rho|<1$. The parameter of interest is β. If $\rho\neq 0$, then y_2 is endogenous in (S16) and ordinary least squares will be inconsistent for β.

Let $M_X\equiv I-X(X^{\mathrm T}X)^{-1}X^{\mathrm T}$ and define $z\equiv M_Xz_1$. The instrumental variables, or two stage least squares, estimator for β is

$$\hat\beta=\beta+\frac{z^{\mathrm T}u}{\pi z^{\mathrm T}z+z^{\mathrm T}v}.$$

The estimator is invariant with respect to γ and δ. Its distribution is largely determined by the so-called concentration parameter $\mu^2\equiv\pi^2z^{\mathrm T}z/\sigma_v^2$, and hence by π. The concentration parameter measures the strength of the instruments. If $\pi \neq 0$ is constant, then the instruments are called strong, and the large T asymptotic distribution of $\hat{\beta}$ is Gaussian, centered around the true parameter. Different asymptotic results are obtained if π is modelled as local to zero. Specifically, if π is $\mathcal{O}(n^{-1/2})$, then the instruments are termed weak, and the asymptotic distribution is that of a ratio of normals (Staiger & Stock, 1997).

Davidson & MacKinnon (2010) discuss several bootstrap schemes for the null distribution of $\hat{\beta}$. We consider what they refer to as the wild restricted efficient residual bootstrap, which they found to perform most favorably. Generating bootstrap samples for y_2 from resampled residuals requires an estimate for π, but no consistent estimator exists if the instruments are weak. Davidson & MacKinnon attempt to mitigate this by using an estimator $\hat{\pi}$ that is asymptotically equivalent to three stage least squares applied to the system (S16) and (S17) under the null, and hence more efficient than the ordinary least squares estimator from the first stage regression (S17). A bootstrap replication is then constructed from the resampled residuals $[\hat{u}\{T/(T-k)\}^{1/2} \odot r, \hat{v}\{T/(T-k-1)\}^{1/2} \odot r]$, where r is a $T \times 1$ vector of independent Rademacher random variables, $\odot$ denotes elementwise multiplication, and $\hat{u}$ and $\hat{v}$ are the residuals from efficient estimation under the null. Davidson & MacKinnon studentize the estimator with a heteroskedasticity consistent standard error. We do not pursue this here and resample the estimator directly. The studentized statistic is not asymptotically pivotal if the instruments are weak (Dufour, 1997), so is not clear that studentization is beneficial. This has been demonstrated by Hirschberg & Lye (2005) for the Fieller–Creasy problem and by Kilian (1999) in a different setting, and is borne out in the Monte Carlo experiment reported below.

Under $H_0 : \beta = \beta_0$, $\hat{\beta} - \beta_0 = X/Y$, where $X = z^{\mathrm{T}}u$ and $Y = \pi z^{\mathrm{T}}z + z^{\mathrm{T}}v$, and the joint bootstrap cumulant generating function of (X, Y) is

$$\mathbb{K}(s,t) = t\hat{\pi}z^{\mathrm{T}}z - T\log(2) + \sum_{j=1}^{T} z_j(s\hat{u}_j + t\hat{v}_j) + \log\left\{1 + e^{-2z_j(s\hat{u}_j + t\hat{v}_j)}\right\}.$$

The sample size T plays a similar role as n in Theorems 3 and 5, but the summands constituting X and Y, while independent, are not identically distributed. Therefore, we apply the saddlepoint approximation formally with $n = 1$. Conditions under which Theorem 4 continues to hold in this case are discussed in Kolassa (2006, Sec. 5·2) for $d = 1$; see also Robinson et al. (1990, Sec. 2·2). The essential requirement is that $\mu^2/T \to C > 0$ as $T \to \infty$, implying strong instruments.

We conducted a small Monte Carlo experiment to assess whether eschewing studentization affects the size of the bootstrap test. The data generating process is the same as Davidson & MacKinnon's. Specifically, the instrument z_1 in (S17) is drawn from a standard Gaussian and normalized to have unit Euclidian norm. The structural innovation in (S16) is generated by the heteroskedastic process $u_t = |z_{1,t}|\varepsilon_{1,t}\sqrt{T}$, and the reduced form error is modeled as $v_t = \rho u_t + (1-\rho^2)^{1/2}\varepsilon_{2,t}$. Here $\varepsilon_{1,t}$ and $\varepsilon_{1,t}$ are independent standard normal. These choices ensure that the concentration parameter $\mu^2 = \pi^2$. An intercept is included as the only exogenous regressor. We fix $\pi \in \{2, 8\}$ and $\rho \in \{0{\cdot}1, 0{\cdot}9\}$ and vary T between 25 and 400, essentially replicating Davidson & MacKinnon's Table 7. The only difference is that we consider an exactly identified model, whereas Davidson & MacKinnon include ten irrelevant instruments along with z_1. As discussed below, our approach does not carry over to that setting. Like Davidson & MacKinnon we use $B = 399$ bootstrap replications, as the sample sizes are too large to compute the exact bootstrap distribution. The $\alpha\%$ equal tail bootstrap p-value is $2\min(f, 1-f)$, where f is the fraction of bootstrapped values of $\hat{\beta}$ not exceeding the observed estimator.

The results of 100,000 Monte Carlo draws are shown in Figure S1. The nominal size is 5%. It is seen that using the saddlepoint approximation instead of sampling from the empirical distribution

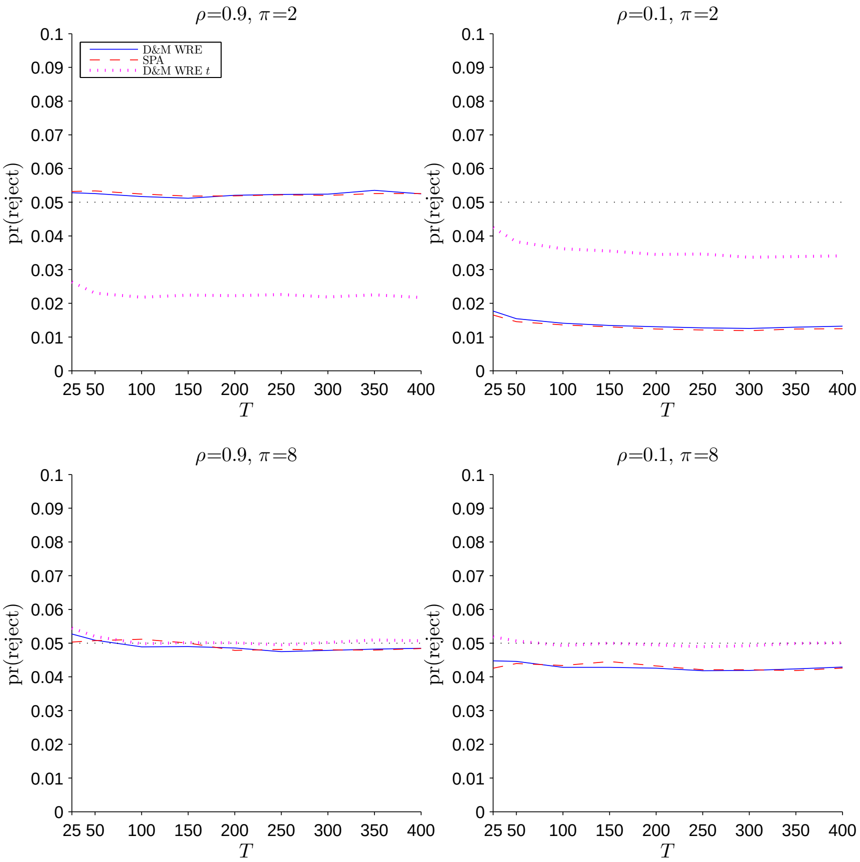

Fig. S1. Sizes of nominal 5% equal tail bootstrap tests of $H_0 : \beta = 0$. Solid, Monte Carlo without studentization; dashes, first order saddlepoint approximation without studentization; dots, Monte Carlo with heteroskedasticity consistent studentization.

function of the residuals makes little difference for the empirical size of the test. Regarding the relative performance of the studentized and unstudentized tests, the former appears to be undersized when the instruments are very weak, and the latter when the endogeneity is weak. Neither is clearly preferred in this experiment.

As mentioned, we briefly discuss the possibility of extending the method to the overidentified case. Let $Z \equiv M_X(z_1\ Z_2)$, where Z_2 has dimension $T \times l$ and l is the degree of overidentification. The two stage least squares estimator for β is now

$$\hat{\beta} = \frac{y_2^{\mathrm{T}} P_Z y_1}{y_2^{\mathrm{T}} P_Z y_2} = \beta + \frac{\pi^{\mathrm{T}} Z^{\mathrm{T}} u + v^{\mathrm{T}} P_Z u}{\pi^{\mathrm{T}} Z^{\mathrm{T}} Z \pi + 2\pi^{\mathrm{T}} Z^{\mathrm{T}} v + v^{\mathrm{T}} P_Z v}, \tag{S18}$$

where $P_Z \equiv Z^{\mathrm{T}}(Z^{\mathrm{T}}Z)^{-1}Z$ and π contains the coefficients on Z in the reduced form equation. The matrix P_Z is positive semidefinite, so that the denominators of the fractions in (S18) are almost surely positive, and the bootstrap distribution could be approximated by the standard result of Daniels (1954) if the joint cumulant generating function of $\pi^{\mathrm{T}} Z^{\mathrm{T}} u + v^{\mathrm{T}} P_Z u$ and $\pi^{\mathrm{T}} Z^{\mathrm{T}} Z \pi + 2\pi^{\mathrm{T}} Z^{\mathrm{T}} v + v^{\mathrm{T}} P_Z v$ were tractable. Unfortunately this is not the case. To see this, consider $E\{\exp(v^{\mathrm{T}} P_Z v)\}$. Unlike in the Gaussian case, one cannot use the spectral theorem to reduce $v^{\mathrm{T}} P_z v$ to a sum of independent random variables. Consequently, computing $E\{\exp(v^{\mathrm{T}} P_Z v)\}$ requires enumerating all 2^T possible realizations for v, which becomes infeasible quickly and renders the use of the approximation moot.

REFERENCES

ANDERSON, T. W. & SAWA, T. (1973). Distributions of estimates of coefficients of a single equation in a simultaneous system and their asymptotic expansions. *Econometrica* **41**, 683–714.

BLEISTEIN, N. (1966). Uniform expansions of integrals with stationary points near algebraic singularity. *Comm. Pure Appl. Math.* **19**, 353–370.

BRODA, S. A. (2013). Tail probabilities and partial moments for quadratic forms in multivariate generalized hyperbolic random vectors. Tinbergen Institute Discussion Paper 13-001/III, Amsterdam, The Netherlands.

BURGESS, S. & THOMPSON, S. G. (2015). *Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation*. Boca Raton: Chapman and Hall/CRC.

DANIELS, H. E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Statist.* **25**, 631–650.

DAVIDSON, R. & MACKINNON, J. G. (2010). Wild bootstrap tests for IV regression. *J. Bus. Econ. Statist.* **28**, 128–144.

DUFOUR, J.-M. (1997). Some impossibility theorems in econometrics with applications to structural and dynamic models. *Econometrica* **65**, 1365–1387.

FIELLER, E. C. (1932). The distribution of the index in a normal bivariate population. *Biometrika* **24**, 428–440.

FORCHINI, G. (2006). On the bimodality of the exact distribution of the TSLS estimator. *Econ. Theory* **22**, 932–946.

FORCHINI, G. (2007). The exact distribution of the TSLS estimator for a non-gaussian just-identified linear structural equation. *Econ. Letters* **95**, 117–123.

FORCHINI, G. & HILLIER, G. (2003). Conditional inference for possibly unidentified structural equations. *Econ. Theory* **19**, 707–743.

FORREY, R. C. (1997). Computing the hypergeometric function. *J. Comput. Phys.* **137**, 79–100.

GORDON, R. D. (1941). Values of Mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *Ann. Math. Statist.* **12**, 364–366.

GRADSHTEYN, I. S. & RYZHIK, I. M. (2007). *Table of Intergrals, Series, and Products*. Burlington: Academic Press, 7th ed.

HILLIER, G. (2006). Yet more on the exact properties of IV estimators. *Econ. Theory* **22**, 913–931.

HINKLEY, D. V. (1969). On the ratio of two correlated normal random variables. *Biometrika* **56**, 635–639.

HIRSCHBERG, J. G. & LYE, J. N. (2005). Inferences for the extremum of quadratic regression models. Department of Economics Research Paper 906, University of Melbourne.

HOLLY, A. & PHILLIPS, P. C. B. (1979). A saddlepoint approximation to the distribution of the k-class estimator of a coefficient in a simultaneous system. *Econometrica* **47**, 1527–1547.

KILIAN, L. (1999). Finite-sample properties of percentile and percentile-t bootstrap confidence intervals for impulse responses. *Rev. Econ. Statist.* **81**, 652–660.

KNIGHT, J. L. (1986). Non-normal errors and the distribution of OLS and 2SLS structural estimators. *Econ. Theory* **2**, 75–106.

KOLASSA, J. E. (2006). *Series Approximation Methods in Statistics*. New York: Springer, 3rd ed.

KOLASSA, J. E. & LI, J. (2010). Multivariate saddlepoint approximations in tail probability and conditional inference. *Bernoulli* **16**, 1191–1207.

LI, J. (2009). Saddle point approximation. Ph.D. Thesis, Rutgers University.

LUGANNANI, R. & RICE, S. O. (1980). Saddlepoint approximations for the distribution of sums of independent random variables. *Adv. in Appl. Probab.* **12**, 475–90.

MADDALA, G. S. & JEONG, J. (1992). On the exact small sample distribution of the instrumental variable estimator. *Econometrica* **60**, 181–183.

NELSON, C. R. & STARTZ, R. (1990a). The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one. *J. Bus.* **63 (Part 2)**, S125–S140.

NELSON, C. R. & STARTZ, R. (1990b). Some further results on the exact small sample properties of the instrumental variable estimator. *Econometrica* **58**, 967–976.

PHILLIPS, P. C. B. (2006). A remark on bimodality and weak instrumentation in structural equation estimation. *Econ. Theory* **22**, 947–960.

RICE, S. O. (1968). Uniform asymptotic expansions for saddle point integrals – applications to a probability distribution occuring in noise theory. *Bell Syst. Tech. J.* **47**, 1971–2013.

RICHARDSON, D. H. (1968). The exact distribution of a structural coefficient estimator. *J. Amer. Statist. Assoc.* **63**, 1214–1225.

ROBINSON, J., HÖGLUND, T., HOLST, L. & QUINE, M. P. (1990). On approximating probabilities for small and large deviations in $\mathbb{R}^d$. *Ann. Probab.* **18**, 727–753.

SAWA, T. (1969). The exact sampling distributions of ordinary least squares and two stage least squares estimates. *J. Amer. Statist. Assoc.* **64**, 923–980.

STAIGER, D. & STOCK, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica* **65**, 557–586.

WANG, S. (1990). Saddlepoint approximations for bivariate distributions. *J. Appl. Probab.* **27**, 586–597.

WOGLOM, G. (2001). More results on the exact small sample properties of the instrumental variable estimator. *Econometrica* **69**, 1381–1389.