Proof of Theorem 2. 1. Let X_n denote the $n \times 2$ design matrix for the Egger regression, i.e., a column of $1's$ and a column of the regressors S_j. Let $\vec{Z}$ denote the column vector of responses $Z_j = Y_j/\sigma_j$. The coefficient $\hat{\beta}_0$, the first component of $(X_n^T X_n)^{-1} X_n^T \vec{Z}$, is $\hat{\beta}_0 = (\hat{\text{Var}}(\vec{S}))^{-1} n^{-1} \sum_{j=1}^n (\overline{S^2} - \overline{S} S_j) Z_j$, where $\hat{\text{Var}}(\vec{S}) = n^{-1} \sum_j (S_j - \overline{S})^2$. The variance estimate is $\hat{\text{Var}}(\hat{\beta}_0) = \frac{SSE \cdot \overline{S^2}}{n(n-2)\hat{\text{Var}}(\vec{S})}$, where $SSE = |(I - X(X^T X)^{-1} X^T)\vec{Z}|_2^2$. The Egger test statistic may therefore be written in terms of Z and S as

$$\frac{\sqrt{n}\hat{\beta}_0}{\sqrt{\hat{\text{Var}}(\hat{\beta}_0)}} = \frac{n^{-1/2} \sum_{j=1}^n (\overline{S^2} - \overline{S} S_j) Z_j}{\sqrt{\overline{S^2} \hat{\text{Var}}(\vec{S}) SSE/(n-2)}}. \tag{1}$$

The numerator in (1) may be replaced by an asymptotically equivalent IID normalized average,

$$\sqrt{n} \sum_{j=1}^n (\overline{S^2} - \overline{S} S_j) Z_j - \sqrt{n} \sum_{j=1}^n (E(S^2) - E(S) S_j) Z_j \underset{P^{(n)}}{\to} 0.$$

By the CLT, $\sqrt{n}\left(E(S^2) - \overline{S^2}\right)$ is $O_{P_n}(1)$. By a weak LLN for trianular arrays $\overline{Z} \to_{P^{(n)}} 0$. Therefore $\sqrt{n}\left(E(S^2) - \overline{S^2}\right)\overline{Z} \to 0$ in probability along $P^{(n)}$. Similarly, the CLT implies $\sqrt{n}\left(E(S) - \overline{S}\right)$ is $O_{P_n}(1)$. Orthogonality of Z and S and uniform integrability of Z under $P^{(n)}$ implies $\overline{SZ} \to_{P_n} E_0(SZ) = 0$. Therefore $\overline{SZ}\left(E(S) - \overline{S}\right) \to 0$ along $P^{(n)}$.

For the denominator, $SSE/(n-2) = \vec{Z}^T (I - X(X^t X)^{-1} X^t)\vec{Z}/n \to \text{Var}(Z)$ in probability along $P^{(n)}$. The form

$$\begin{aligned}\zeta_n^T X(X^T X)^{-1} X^T \vec{Z} &= (\hat{\text{Var}}(\vec{S}))^{-1}\left(\overline{Z}(\overline{S^2}\ \overline{Z} - \overline{S}\ \overline{SZ}) + \overline{ZS}(\overline{ZS} - \overline{Z}\ \overline{S})\right) \\ &= (\hat{\text{Var}}(\vec{S}))^{-1}\left(\overline{S^2}(\overline{Z})^2 + (\overline{ZS})^2 - 2\overline{S}\ \overline{Z}\ \overline{SZ}\right).\end{aligned}$$

converges in probability to 0 along $P^{(n)}$, with each monomial converging to $E(S^2)E(Z^2)$. Therefore $SSE/(n-2) = n^{-1}\vec{Z}^T\vec{Z} + o_{P^{(n)}}(1)$, which tends along $P^{(n)}$ to $E(Z^2) = \text{Var}(Z)$, as above.

2. By part 1, the standardized test statistic is asymptotically equivalent to an IID average, and asymptotic normality follows from the Lindeberg-Feller CLT. The conditions of that theorem follow from the assumptions here on S and Z, which in fact imply that the random variables $\{\text{Var}(S)^{-1}(E(S^2) - E(S)S_1)Z_1 - E_n(Z)\}$, as the distribution P_n varies, are L^2-bounded.

The local limiting power at the null $\theta = 0$ is then

$$\lim_n P^{(n)}\left(\frac{\hat\beta_0}{\sqrt{\operatorname{Var}(\hat\beta_0)}} > t_{n-1,1-\alpha}\right) = \lim_n P^{(n)}\left(n^{-1/2}\frac{\sum_{j=1}^n \operatorname{Var}(S)^{-1}(E(S^2) - E(S)S_j)Z_j}{\sqrt{\operatorname{Var}(Z)E(S^2)/\operatorname{Var}(S)}} > t_{n-1,1-\alpha}\right)$$

$$= \lim_n P^{(n)}\left(n^{-1/2}\frac{\sum_{j=1}^n (\operatorname{Var}(S)^{-1}(E(S^2) - E(S)S_j)Z_j - \mu_n)}{\sqrt{\operatorname{Var}(Z)E(S^2)/\operatorname{Var}(S)}} > t_{n-1,1-\alpha} - \frac{n^{-1/2}\mu_n}{\sqrt{\operatorname{Var}(Z)E(S^2)/\operatorname{Var}(S)}}\right)$$

$$= 1 - \Phi\left(z_{1-\alpha} - \frac{h}{\sqrt{\operatorname{Var}(Z)E(S^2)/\operatorname{Var}(S)}}\right).$$

□

Proof of Theorem 3. Decomposing the left hand side of (3) as

$$\sqrt{n}\left(\hat\tau(\hat\theta) - (E^{(n)}\hat\tau)\left(\frac{\mu_1}{\mu_2}\theta^{(n)}\right)\right) = \sqrt{n}\left(\hat\tau(\hat\theta) - (\Pi^{(n)}\hat\tau)(\hat\theta)\right) \quad (2)$$

$$+ \sqrt{n}((\Pi^{(n)}\hat\tau)(\hat\theta) - (E^{(n)}\hat\tau)(\hat\theta)) \quad (3)$$

$$+ \sqrt{n}\left((E^{(n)}\hat\tau)(\hat\theta) - (E^{(n)}\hat\tau)\left(\frac{\mu_1}{\mu_2}\theta^{(n)}\right)\right), \quad (4)$$

the proof is divided in 3 steps:

1. $\sqrt{n}\left(\hat\tau(\hat\theta) - (\Pi^{(n)}\hat\tau)(\hat\theta)\right) = o_{\mathbb{P}^{(n)}}(1)$
2. $\sqrt{n}((\Pi^{(n)}\hat\tau)(\hat\theta) - (E^{(n)}\hat\tau)(\hat\theta)) = \sqrt{n}(\Pi^{(n)}\hat\tau)(0) + o_{\mathbb{P}^{(n)}}(1)$
3. $\sqrt{n}\left((E^{(n)}\hat\tau)(\hat\theta) - (E^{(n)}\hat\tau)\left(\frac{\mu_1}{\mu_2}\theta^{(n)}\right)\right) = \sqrt{n}\left(\frac{\overline{ZS}}{\mu_2} - \frac{\mu_1}{\mu_2}\theta^{(n)}\right) \cdot 2E|S_1 - S_2|E^{(\infty)}f_Z^{(\infty)}(Z) + o_{\mathbb{P}^{(n)}}(1)$.

1. Let

$$g^{(n)}(\theta) = \sqrt{n}\left(\hat\tau(\theta) - \Pi^{(n)}\hat\tau(\theta)\right),$$

so the right-hand side of (2) is $g^{(n)}(\hat\theta)$. So it is enough to show that $\sup_\theta |g^{(n)}(\theta)| = O_{P^{(n)}}(n^{-1/2})$.

$$g^{(n)}(\theta) = \sqrt{n}\left(\hat\tau(\theta) - \Pi^{(n)}\hat\tau(\theta)\right)$$
$$= 2\sqrt{n}\binom{n}{2}^{-1}\sum_{1\le j<k\le n}\left(\left\{\frac{Z_j - Z_k}{S_j - S_k} < \theta\right\} - P^{(n)}\left(\frac{Z_j - Z_k}{S_j - S_k} < \theta \,\middle|\, Z_j, S_j\right)\right.$$
$$\left. - P^{(n)}\left(\frac{Z_j - Z_k}{S_j - S_k} < \theta \,\middle|\, Z_k, S_k\right) + P^{(n)} \otimes P^{(n)}\left(\frac{Z_j - Z_k}{S_j - S_k} < \theta\right)\right).$$

For any fixed θ and n, the last expression is $2\sqrt{n}$ times a U-statistic with bivariate kernel

$$((z,s),(z',s')) \mapsto \left\{\frac{z-z'}{s-s'} < \theta\right\} - P^{(n)}\left(\frac{z-z'}{s-s'} < \theta \middle| z, s\right) \\ - P^{(n)}\left(\frac{z-z'}{s-s'} < \theta \middle| z', s'\right) + P^{(n)} \otimes P^{(n)}\left(\frac{z-z'}{s-s'} < \theta\right). \tag{5}$$

Let $\tilde{\mathcal{F}}^{(n)}$ denote the class of functions (5) as $\theta \in \mathbb{R}$ varies. The kernels in this class give rise to degenerate U-statistics with respect to $\mathbb{P}^{(n)}$, i.e., for any $f \in \tilde{\mathcal{F}}^{(n)}$ and (z,s),

$$P^{(n)} f((Z,S),(Z',S') \mid (Z,S) = (z,s)) = 0.$$

Nolan and Pollard (1987, 1988) provides a bound for the supremum of $|g^{(n)}(\theta)|$ over a class of degenerate bivariate kernels such as $\tilde{\mathcal{F}}^{(n)}$. Given an IID sample $x_1, \ldots, x_{2n}$, let T_n denote the random measure that places mass 1 on all pairs of observations of the form $(x_{2j}, x_{2k}), (x_{2j-1}, x_{2k-1}), (x_{2j-1}, x_{2k})$, or (x_{2j}, x_{2k-1}), where $j \neq k, 1 \leq j, k \leq n$. Given a class of real function $\mathcal{F}$, measure μ, and $u > 0$, let $N(u, \mu, \mathcal{F})$ denote the $L^2(\mu)$ covering number of $\mathcal{F}$, and let $J(\mu, \mathcal{F}) = \int_0^1 \log N(u, \mu, \mathcal{F}) du$ denote the associated covering integral. Let F denote a bound for the functions in $\tilde{\mathcal{F}}^{(n)}$. By Theorem 6 of Nolan and Pollard (1987), there is a constant c such that

$$\mathbb{P}^{(n)} \sup_{\tilde{\mathcal{F}}^{(n)}} |g^{(n)}| \leq \frac{c}{\sqrt{n}} \mathbb{P}^{(n)} \left(\frac{1}{4} \sup_{\tilde{\mathcal{F}}^{(n)}} \sqrt{\frac{T_n f^2}{n^2}} + \sqrt{\frac{T_n F^2}{n^2}} J\left(T_n, \tilde{\mathcal{F}}^{(n)}\right)\right).$$

With two applications of the Cauchy-Schwarz inequality,

$$\mathbb{P}^{(n)} \sup_{\tilde{\mathcal{F}}^{(n)}} |g^{(n)}| \leq \frac{c}{\sqrt{n}} \sqrt{\mathbb{P}^{(n)} \frac{T_n F^2}{n^2}} \left(1 + \sqrt{\mathbb{P}^{(n)} J(T_n, \tilde{\mathcal{F}}^{(n)})^2}\right).$$

Taking the bound $F = 4$ for $f \in \tilde{\mathcal{F}}^{(n)}$, $\sqrt{\mathbb{P}^{(n)} \frac{T_n F^2}{n^2}} \leq 64$. Since $\tilde{\mathcal{F}}^{(n)} \subset \mathcal{F} + 2P^{(n)}\mathcal{F} + P^{(n)} \otimes P^{(n)}\mathcal{F}$, Lemma 16 of Nolan and Pollard (1987), implies

$$N(u, T_n, \tilde{\mathcal{F}}^{(n)}) \leq N(u/4, T_n, \mathcal{F}) \cdot N(u/16, T_n, P^{(n)}\mathcal{F}) \cdot N(u/64, T_n, P^{(n)}\mathcal{F}) \cdot N(u/64, T_n, P^{(n)} \otimes P^{(n)}\mathcal{F}).$$

The functions in each of the classes $\mathcal{F}, P^{(n)}\mathcal{F}$, and $P^{(n)} \otimes P^{(n)}\mathcal{F}$ are monotonic in θ, so each has a linear discriminating polynomial, $p(x) = x + 1$. By the Approximation Lemma, II.25 of Pollard (1984), there exist constants A, W, depending only on the discriminating polynomial, such that

$$N(u, \tilde{\mathcal{F}}^{(n)}) \leq A\left(\frac{u}{4}\right)^{-W} \cdot A\left(\frac{u}{16}\right)^{-W} \cdot 2A\left(\frac{u}{64}\right)^{-W} = 2A^3\left(\frac{u^3}{4^6}\right)^{-W}.$$

Therefore,

$$J(T_n, \tilde{\mathcal{F}}^{(n)}) = \int_0^1 \log N(u, T_n, \tilde{\mathcal{F}}^{(n)}) du \leq \int_0^1 \log\left(2A^3\left(\frac{u^3}{4^6}\right)^{-W}\right) du$$

is bounded uniformly in n.

2. Let $P_n^{(n)}$ denote the empirical measure on a sample of size n under $P^{(n)}$. For $\theta \in \mathbb{R}$, let $h_\theta^{(n)}$ denote the function $(z,s) \mapsto 4P^{(n)}\left(\frac{Z-Z'}{S-S'} < \theta \middle| Z=z, S=s\right)$. Then

$$\sqrt{n}\left((\Pi^{(n)}\hat{\tau})(\theta) - (E^{(n)}\hat{\tau})(\theta)\right) = \sqrt{n}\left(\frac{4}{n}\sum_{j=1}^{n} P^{(n)}\left(\frac{Z_j - Z}{S_j - S} < \theta \middle| Z_j, S_j\right) - 4P^{(n)} \otimes P^{(n)}\left(\frac{Z-Z'}{S-S'} < \theta\right)\right)$$

$$= \sqrt{n}\left(P_n^{(n)} - P^{(n)}\right)(h_\theta^{(n)}).$$

For fixed n, letting θ vary, $\sqrt{n}\left((\Pi^{(n)}\hat{\tau})(\theta) - (E^{(n)}\hat{\tau})(\theta)\right)$ is therefore an empirical process indexed by the class of functions

$$P^{(n)}\mathcal{F} = \left\{h_\theta^{(n)} : \theta \in \mathbb{R}\right\}.$$

Letting n vary gives a sequence of processes, in terms of which (3) is

$$\sqrt{n}\left(P_n^{(n)} - P^{(n)}\right)(h_{\hat{\theta}}^{(n)} - h_0^{(n)}) = o_{\mathbb{P}^{(n)}}(1),$$

which follows on showing

(a) $\sqrt{n}\left(P_n^{(n)} - P^{(n)}\right)$ is stochastically equicontinuous along $P^{(n)}$, that is, for any $\eta > 0, \epsilon > 0$, there is $\delta > 0$ such that

$$\limsup_{n\to\infty} \mathbb{P}^{(n)}\left(\sup_{[\delta]^{(n)}} |\sqrt{n}\left(P_n^{(n)} - P^{(n)}\right)(h_\theta^{(n)} - h_{\theta'}^{(n)})| > \eta\right) < \epsilon,$$

where $[\delta]^{(n)} = \{(\theta,\theta') \in P^{(n)}\mathcal{F}^{(n)} : P^{(n)}(h_\theta^{(n)} - h_{\theta'}^{(n)})^2 \leq \delta^2\}$,

(b) For any $\epsilon > 0$, there is $\delta > 0$ such that $|\theta| < \delta$ implies $\limsup P^{(n)}(h_\theta^{(n)} - h_0^{(n)})^2 < \epsilon$, and

(c) $\hat{\theta}$ is consistent for 0 along $P^{(n)}$, i.e., for any $\delta > 0, \epsilon > 0$, $\limsup P^{(n)}(|\hat{\theta}| > \delta) < \epsilon$.

(a) Were $P^{(n)}$ and $\mathcal{F}^{(n)}$ fixed, the stochastic equicontinuity of $\sqrt{n}\left(P_n^{(n)} - P^{(n)}\right)$ would follow from standard empirical process theory. A small variation of the Equicontinuity Lemma, Theorem VII.15 of Pollard (1984), accommodates changing probability measures and function classes. The variation presented below ignores measurability qualifications that are not relevant for the present application.

Lemma 1. *Given function classes $\mathscr{F}^{(n)}$ and probability measures $\mathscr{P}^{(n)}, n \in \mathbb{N}$. Assume for any $\epsilon > 0, \eta > 0$, there is $\gamma > 0$ such that*

$$\limsup_n \mathscr{P}^{(n)}\left(J\left(\gamma, \mathscr{P}^{(n)}, \mathscr{F}^{(n)}\right) > \eta\right) < \epsilon.$$

Then there exists $\delta > 0$ such that

$$\limsup_n \mathscr{P}^{(n)}\left(\sup_{[\delta]^{(n)}} |\left(\mathscr{P}_n^{(n)} - \mathscr{P}^{(n)}\right)(f-g)| > \eta\right) < \epsilon,$$

where $[\delta]^{(n)} = \left\{(f,g) \in \mathscr{F}^{(n)} : \mathscr{P}^{(n)}(f-g)^2 < \delta^2\right\}$.

The proof follows by superficial changes to the proof of the form for fixed $\mathscr{F}^{(n)}$ and $\mathscr{P}^{(n)}$ cited above, and is omitted.

The stochastic equicontinuity (2a) follows from the conclusion of the Lemma by setting $\mathscr{P}^{(n)} = P^{(n)}, \mathscr{F}^{(n)} = \mathcal{P}^{(n)}\mathcal{F}$. The assumptions of the Lemma hold by similar arguments as in 1. That is, the functions in $\mathcal{F}^{(n)}$ are monotonic in θ, so the graphs have discriminating polynomial $p(x) = x + 1$ for all n, and then it follows from the Approximation Lemma, II.25 of Pollard (1984), that $J\left(\gamma, P^{(n)}, \mathcal{F}^{(n)}\right)$ is $O(\gamma)$ deterministically.

(b) For $\theta > 0$,

$$\begin{aligned}
P^{(n)}(h_\theta^{(n)} - h_0^{(n)})^2 &= 16P^{(n)}\left(P^{(n)}\left(0 < \frac{Z - Z'}{S - S'} < \theta \,\middle|\, Z, S\right)\right)^2 \\
&\leq 16P^{(n)}\left(0 < \frac{Z - Z'}{S - S'} < \theta\right) \\
&= 16E^{(n)}\left(F_{Z-Z'}^{(n)}(\theta|S - S'|) - 1/2\right).
\end{aligned}$$

For any $\epsilon > 0$ and large enough n, $|F^{(n)} - F^{(\infty)}|_\infty < \epsilon/32$, so $F_{Z-Z'}^{(n)}(\theta|S - S'|) - 1/2 < F_{Z-Z'}^{(\infty)}(\theta|S - S'|) - 1/2 + \epsilon/32 = \int_0^{\delta|S-S'|} f_{Z-Z'}^{(\infty)}(u)du + \epsilon/32$. By the Dominated Convergence Theorem, and since the law of S doesn't change with n, there is $\delta > 0$ such that for large enough n, $E^{(n)} \int_0^{\delta|S-S'|} f_{Z-Z'}^{(\infty)}(u)du < \epsilon/32$. For such n, the last expression of the above display is $\leq \epsilon$.

(c) Asymptotic normality is established in 3 below, implying consistency.

3. Expanding (4) to first order,

$$\begin{aligned}
&\sqrt{n}\left(\left(E^{(n)}\hat{\tau}\right)(\hat{\theta}) - \left(E^{(n)}\hat{\tau}\right)\left(\frac{\mu_1}{\mu_2}\theta^{(n)}\right)\right) \\
&\quad = \sqrt{n}\left(\hat{\theta} - \frac{\mu_1}{\mu_2}\theta^{(n)}\right)\frac{\mathrm{d}}{\mathrm{d}\theta}\left(E^{(n)}\hat{\tau}\right)(\theta)\bigg|_{\theta=\frac{\mu_1}{\mu_2}\theta^{(n)}} + \sqrt{n}\cdot o\left(\hat{\theta} - \frac{\mu_1}{\mu_2}\theta^{(n)}\right).
\end{aligned}$$

As $\left(E^{(n)}\hat{\tau}\right)(\theta) = 2\mathbb{P}^{(n)}\left(\frac{Z-Z'}{S-S'} < \theta\right) - 1 = E\left(2F_{Z-Z'}^{(n)}(\theta|S - S'|) - 1\right)$, the derivative of $\theta \mapsto \left(E^{(n)}\hat{\tau}\right)(\theta)$ is

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\theta}\left(E^{(n)}\hat{\tau}\right)(\theta) &= \frac{\mathrm{d}}{\mathrm{d}\theta}E\left(2F_{Z-Z'}^{(n)}(\theta|S - S'|) - 1\right) \\
&= E\left(2|S - S'|f_{Z-Z'}^{(n)}(\theta|S - S'|)\right) \\
&= 2f_{Z-Z'}^{(\infty)}(0)E|S - S'| + 2E\left(|S - S'|\left(f_{Z-Z'}^{(\infty)}(\theta|S - S'|) - f_{Z-Z'}^{(\infty)}(0)\right)\right) + o(1).
\end{aligned} \tag{6}$$

The derivative may be brought inside the expectation in the second equality since the derivative $\frac{\mathrm{d}}{\mathrm{d}\theta}\left(2F_{Z-Z'}^{(n)}(\theta|S - S'|) - 1\right) = 2|S - S'|f_{Z-Z'}^{(n)}(\theta|S - S'|)$ is nonnegative. Furthermore, the error $2E\left(|S-S'|\left(f_{Z-Z'}^{(\infty)}(\theta|S-S'|)-f_{Z-Z'}^{(\infty)}(0)\right)\right) \to 0$ as $\theta \to 0$ by dominated convergence, since model 14 implies $S-S' \in L^1$, and assumption 2, $f^{(\infty)} \in L^2$, implies $f^{(\infty)} * f^{(\infty)}$ is bounded.

Substituting the derivative expression into (6),

$$
\begin{aligned}
&\sqrt{n}\left(\left(E^{(n)}\hat{\tau}\right)(\hat{\theta})-\left(E^{(n)}\hat{\tau}\right)\left(\frac{\mu_1}{\mu_2}\theta^{(n)}\right)\right) \\
&\quad=\sqrt{n}\left(\hat{\theta}-\frac{\mu_1}{\mu_2}\theta^{(n)}\right)\cdot 2E|S_1-S_2|E^{(\infty)}f_Z^{(\infty)}(Z)+\sqrt{n}\cdot o\left(\hat{\theta}-\frac{\mu_1}{\mu_2}\theta^{(n)}\right) \\
&\quad=\sqrt{n}\left(\hat{\theta}-\frac{\mu_1}{\mu_2}\theta^{(n)}\right)\cdot 2E|S_1-S_2|E^{(\infty)}f_Z^{(\infty)}(Z)+\sqrt{n}\left(\hat{\theta}-\frac{\mu_1}{\mu_2}\theta^{(n)}\right)\cdot o(1)+\sqrt{n}\cdot o\left(\hat{\theta}-\frac{\mu_1}{\mu_2}\theta^{(n)}\right).
\end{aligned}
$$

Finally, $\sqrt{n}\left(\hat{\theta}-\frac{\mu_1}{\mu_2}\theta^{(n)}\right)$ is asymptotically equivalent under $P^{(n)}$ to a centered and scaled IID average,

$$
\begin{aligned}
\sqrt{n}\left(\hat{\theta}-\frac{\mu_1}{\mu_2}\theta^{(n)}\right) &= \sqrt{n}\left(\frac{\sum Z_jS_j}{\sum S_j^2}-\frac{\mu_1}{\mu_2}\theta^{(n)}\right) \\
&= \left(\frac{n}{\sum S_j^2}-\frac{1}{\mu_2}\right)\cdot\frac{1}{\sqrt{n}}\sum Z_jS_j+\frac{1}{\sqrt{n}\mu_2}\sum\left(Z_jS_j-\mu_1\theta\right) \\
&= o_{P^{(n)}}(1)\cdot O_{P^{(n)}}(1)+\frac{1}{\sqrt{n}\mu_2}\sum\left(Z_jS_j-\mu_1\theta\right),
\end{aligned}
$$

using again the second moment assumptions 2 and 14.

Therefore $\sqrt{n}\left(\hat{\theta}-\frac{\mu_1}{\mu_2}\theta^{(n)}\right) = \sqrt{n}\left(\frac{1}{\mu_2 n}\sum_j Z_jS_j-\frac{\mu_1}{\mu_2}\theta^{(n)}\right)+o_{\mathbb{P}^{(n)}}(1)$, and the common distributional limit of the left and right hand sides is given by a traingular array CLT,

$$
\begin{aligned}
E^{(n)}\frac{1}{\mu_2 n}\sum_j Z_jS_j &= \theta^{(n)}\frac{\mu_1}{\mu_2} \\
\mathrm{Var}^{(n)}\frac{1}{\mu_2 n}\sum_j Z_jS_j &= \frac{\mathrm{Var}^{(n)}Z}{\mu_2}+\left(\theta^{(n)}\right)^2\left(\frac{1}{\mu_2}-\left(\frac{\mu_1}{\mu_2}\right)^2\right) \\
\sqrt{n}\left(\frac{1}{\mu_2 n}\sum_j Z_jS_j-\frac{\mu_1}{\mu_2}\theta^{(n)}\right) &\overset{\mathbb{P}^{(n)}}{\rightsquigarrow} \mathcal{N}\left(0,\frac{\mathrm{Var}^{(\infty)}Z}{\mu_2}\right)
\end{aligned}
$$

□

Proof of Corollary 4. The proof of Theorem 3 shows that the right-hand side of (2) is uniformly $O_{P^{(n)}}(n^{-1/2})$ and does not contribute to the asymptotic variance. The variance of (3) is 4/9, the approximation used in the uncorrected Begg test, as discussed in Michael and Ghebremichael (2023). By a triangular array LLN the asymptotic variance of (4) is $\operatorname{Var} Z/\mu_2 \left(2E|S_1 - S_2| E f_Z(Z)\right)^2$. The covariance of (3) and (4) is

$$\begin{aligned}
&\operatorname{Cov}\left(\sqrt{n}(\Pi^{(n)}\hat{\tau})(0), \sqrt{n}\left(\frac{\overline{ZS}}{\mu_2} - \frac{\mu_1}{\mu_2}\theta^{(n)}\right) \cdot 2E|S_1 - S_2| E^{(\infty)} f_Z^{(\infty)}(Z)\right) + o(1) \\
&= 2E|S_1 - S_2| E^{(n)} f_Z(Z) \cdot n\operatorname{Cov}^{(n)}(\Pi^{(n)}\hat{\tau}(0), \overline{ZS}/\mu_2) + o(1).
\end{aligned}$$

Using the relation

$$\begin{aligned}
&E^{(n)}\left(Z_1 S_1 \{Z_1 - Z_2 < 0\}\{S_1 - S_2 > 0\}\right) + E^{(n)}\left(Z_1 S_1 \{Z_1 - Z_2 > 0\}\{S_1 - S_2 < 0\}\right) \\
&= -\frac{1}{2} E^{(n)}|S_1 - S_2| E^{(n)}(Z F_Z(Z)),
\end{aligned}$$

it follows that

$$\begin{aligned}
n\operatorname{Cov}^{(n)}(\Pi^{(n)}\hat{\tau}(0), \overline{ZS}/\mu_2) &= n\operatorname{Cov}^{(n)}\left(\frac{2}{n}\sum_{j=1}^{n} 2P^{(n)}\left(\frac{Z_j - Z}{S_j - S} < 0 \middle| Z_j, S_j\right), \frac{1}{n\mu_2}\sum_{j=1}^{n} Z_j S_j\right) \\
&= \frac{4}{\mu_2}\operatorname{Cov}\left(P\left(\frac{Z_1 - Z}{S_1 - S} < 0 \middle| Z_1, S_1\right), Z_1 S_1\right) + o(1) \\
&= \frac{-2}{\mu_2} E|S_1 - S_2| E(Z F_Z(Z)) + o(1).
\end{aligned}$$

□

Proof of Corollary 5. The centering for the test statistic under $P^{(n)}$ used in Theorem 3 is

$$\begin{aligned}
E^{(n)}\hat{\tau}\left(\frac{\mu_1}{\mu_2}\theta^{(n)}\right) &= 2P^{(n)}\left(\frac{Z_1 - Z}{S_1 - S} < \frac{\mu_1}{\mu_2}\theta^{(n)}\right) - 1 \\
&= 2P^{(n)}\left(\left(Z_1 - Z - \frac{\mu_1}{\mu_2}\theta^{(n)}(S_1 - S)\right)(S_1 - S) < 0\right) - 1 \\
&= 4P^{(n)}\left(\left\{Z_1 - Z - \frac{\mu_1}{\mu_2}\theta^{(n)}(S_1 - S) < 0\right\}\{S_1 - S > 0\}\right) - 1.
\end{aligned}$$

The derivative of the probability in the last line at $\lim \theta^{(n)} = 0$ is

$$\begin{aligned}
\frac{d}{d\theta^{(n)}}\int_0^\infty \int_{-\infty}^{\mu_1/\mu_2\theta^{(n)}v} f^{(n)}_{Z_1 - Z_2}(u) f_{S_1 - S_2}(v)\,du\,dv\Big|_{\theta^{(n)}=0} &= \frac{\mu_1}{\mu_2}\int_0^\infty v f^{(n)}_{Z_1 - Z_2}(\mu_1/\mu_2\theta^{(n)}v) f_{S_1 - S_2}(v)\,du\,dv\Big|_{\theta^{(n)}=0} \\
&= f^{(n)}_{Z_1 - Z_2}(0) E(S_1 - S_2; S_1 - S_2 > 0) \\
&= E^{(n)} f^{(n)}_{Z_1 - Z_2}(Z) E|S_1 - S_2|/2.
\end{aligned}$$

Therefore

$$\frac{d}{d\theta^{(n)}} E^{(n)}\hat{\tau}\left(\frac{\mu_1}{\mu_2}\theta^{(n)}\right)\Big|_{\theta^{(n)}=0} \to 2E f_Z(Z) E|S_1 - S_2|.$$

Combining the above expressions for the centering with the expression for the asymptotic variance of $\hat{\tau}(\hat{\theta})$ given in Theorem 3,

$$P^{(n)}\left(\frac{\sqrt{n}\hat{\tau}(\hat{\theta})}{\sqrt{\mathrm{Var}^{(n)}(\hat{\tau}(\hat{\theta}))}} > z_{1-\alpha}\right) = P^{(n)}\left(\sqrt{n}\frac{\hat{\tau}(\hat{\theta}) - (E^{(n)}\hat{\tau})(\mu_1/\mu_2\theta^{(n)})}{\sqrt{\mathrm{Var}^{(n)}(\hat{\tau}(\hat{\theta}))}} > z_{1-\alpha} - \sqrt{n}\frac{(E^{(n)}\hat{\tau})(\mu_1/\mu_2\theta^{(n)})}{\sqrt{\mathrm{Var}^{(n)}(\hat{\tau}(\hat{\theta}))}}\right)$$

$$= P^{(n)}\left(\sqrt{n}\frac{\hat{\tau}(\hat{\theta}) - (E^{(n)}\hat{\tau})(\mu_1/\mu_2\theta^{(n)})}{\sqrt{\mathrm{Var}^{(n)}(\hat{\tau}(\hat{\theta}))}} > z_{1-\alpha} - \sqrt{n}\frac{\mu_1/\mu_2\theta^{(n)}E^{(n)}\hat{\tau})(0)}{\sqrt{\mathrm{Var}^{(n)}(\hat{\tau}(\hat{\theta}))}} + o(1)\right)$$

$$\to 1 - \Phi\left(z_{1-\alpha} - \frac{2\frac{\mu_1}{\mu_2}Ef_Z(Z)E|S_1 - S_2|h}{\sqrt{\frac{4}{9} + 4\frac{(E|S_1-S_2|)^2}{ES^2}Ef_Z(Z)(Ef_Z(Z)\,\mathrm{Var}\,Z - 2E(ZF_Z(Z)))}}\right).$$

□

Proof of Lemma 7. A similar maximization problem was encountered in Michael (2024), with the objective there being instead $E|X_1 - X_2|/\sqrt{EX_1^2}$. As there we use the variational calculus to carry it out.

Maximizing $E|X_1 - X_2|/\sqrt{\mathrm{Var}\,X_1}$ is the same as maximizing $E|X_1 - X_2|$ subject to the constraint that $\mathrm{Var}(X_1) = 1$. Since the objective is unaffected by the mean of the X_i, we take it to be 0. Let F denote the CDF of the X_i, and suppose first that $X \geq x_0$ for an arbitrary number $x_0 < 0$. Michael and Ghebremichael (2023, Theorem 4) gives the representation

$$E|X_1 - X_2| = 2\int_{-\infty}^{\infty} F(x)(1 - F(x))dx, \tag{7}$$

and also shows that the mapping $F \mapsto \int F(x)(1 - F(x))dx$ is concave. The problem can then be stated in terms of the CDF F of X as

$$\begin{aligned}
\text{maximize } & 2\int_{x_0}^{\infty} F(x)(1 - F(x))dx \\
\text{subject to: } & 2\int_{x_0}^{\infty} x(1 - F(x))dx = 1 \\
& \int_{x_0}^{\infty} (1 - F(x))dx = |x_0| \\
& F = 0 \text{ on } \{x < x_0\} \\
& \lim_{x\to\infty} F(x) = 1 \\
& F \text{ monotone}
\end{aligned}$$

The Euler-Lagrange equation is $F(x) = \frac{1}{2}(1 - 2\lambda_1 x - \lambda_2)$, with λ_1, λ_2 to be determined. The monotonicity constraint implies $\lambda_1 \leq 0$ and the constraint $0 \leq F \leq 1$ implies

$$\frac{1 - \lambda_2}{2\lambda_1} < x < -\frac{1 + \lambda_2}{2\lambda_1}.$$

The moment constraints imply $\lambda_1 = -1/\sqrt{12}, \lambda_2 = 0$, regardless of the postulated left support point x_0. Therefore,

$$F(x) = \frac{1}{2} + \frac{1}{2\sqrt{3}}x, \quad -\sqrt{3} \leq x \leq \sqrt{3},$$

which is the CDF of a uniformly distributed RV with mean 0 and variance 1. Applying formula (7), $E|X_1 - X_2| = 2/\sqrt{3}$. □

References

Michael, H. (2024). The effect of screening for publication bias on the outcomes of meta-analyses. *Forthcoming in Scandinavian Journal of Statistics*.

Michael, H. and M. Ghebremichael (2023). A correction to Begg's test for publication bias. *Communications in Statistics - Theory and Methods 0*(0), 1–21.

Nolan, D. and D. Pollard (1987). U-processes: rates of convergence. *The Annals of Statistics*, 780–799.

Nolan, D. and D. Pollard (1988). Functional Limit Theorems for U-Processes. *The Annals of Probability 16*(3), 1291–1298.

Pollard, D. (1984). *Convergence of stochastic processes*. Springer-Verlag New York.