DEPARTMENT OF ECONOMICS AND FINANCE

SCHOOL OF BUSINESS AND ECONOMICS

UNIVERSITY OF CANTERBURY

CHRISTCHURCH, NEW ZEALAND

Meta-Analysis and Partial Correlation Coefficients: A Matter of Weights

Sanghyun Hong
W. Robert Reed

WORKING PAPER

No. 7/2023

Department of Economics and Finance
UC Business School
University of Canterbury
Private Bag 4800, Christchurch
New Zealand

WORKING PAPER No. 7/2023

Meta-Analysis and Partial Correlation Coefficients: A Matter of Weights

Sanghyun Hong[1]
W. Robert Reed[1†]

May 2023

Abstract: This study builds on the simulation framework of a recent paper by Stanley & Doucouliagos (2023). S&D use simulations to make the argument that meta-analyses using partial correlation coefficients (PCCs) should employ a suboptimal estimator of the PCC standard error when constructing weights for fixed effect and random effects estimation. We address concerns that their simulations and subsequent recommendation may give meta-analysts a misleading impression. While the estimator they promote dominates the "correct" formula in their Monte Carlo framework, there are other estimators that perform even better. In addition, their simulations ignore how standard errors are used to test for publication bias. We show how the standard error estimator they recommend gives unsatisfactory results when used in Egger regressions/FAT-PET analyses. We conclude that more research is needed before best practice recommendations can be made for meta-analyses with PCCs.

JEL Classification: C4

Keywords: Partial correlation coefficients, Meta-analysis, Bias, Mean square errors

Data Availability Policy: All the codes used to generate the results in this paper are posted at https://osf.io/zytf3/.

Acknowledgements: Research support (access to supercomputer) from the UC Research Compute Cluster (RCC) is gratefully acknowledged. We are especially thankful to Robbie van Aert, Shinichi Nakagawa, and Yefeng Yang for very helpful comments on an earlier draft of this paper. Remaining errors are ours alone. This research was supported by the Czech Science Foundation (GACR).

[1] Department of Economics and Finance & UCMeta, University of Canterbury, NEW ZEALAND

† Corresponding author: W. Robert Reed. Email: bob.reed@canterbury.ac.nz

1. Introduction

In a recent paper, Stanley and Doucouliagos[1], henceforth S&D, argue that meta-analysts should never use correct standard errors when performing meta-analyses with partial correlation coefficients (PCCs). They present simulations that demonstrate that an alternative, "suboptimal" estimator of the standard error – commonly used in the economics meta-analysis literature – statistically dominates the "correct" estimator when using either random effects, fixed effects, or unrestricted weighted least squares. They recommend its use when the meta-analysis sample is relatively large and the population value of PCC is relatively small.

In this paper, we reproduce S&D's results but argue that their simulations and recommendation may give meta-analysts a misleading impression. We show that S&D's "suboptimal" estimator of the PCC standard error is itself often dominated by other estimators. In S&D's simulation environment, OLS is unbiased, produces reliable confidence intervals and is more efficient than their recommended estimator. Furthermore, their recommended estimator produces unsatisfactory outcomes when used to test for publication bias.

We proceed as follows. Section II describes the research design for S&D's Monte Carlo experiments. Section III demonstrates that we are able to reproduce their results. Section IV notes that the data generating process (DGP) in S&D's Monte Carlo experiments assumes homoskedasticity and effect homogeneity. In this setting, OLS is the optimal estimator and we show that its performance dominates S&D's recommended estimator.

Section V extends their simulation framework to allow for heteroskedasticity and effect heterogeneity. It demonstrates that there are multiple estimators that are superior to S&D's recommended estimator and explains why. Section VI repurposes S&D's simulation environment to show that their recommended estimator produces unsatisfactory outcomes when used in Egger regressions and FAT-PET analyses. Section VII concludes by arguing that there is insufficient understanding of the PCC problem to support a best practice recommendation for meta-analyses with PCCs. Further research is needed.

II. S&D's Research Design

S&D's research design consists of two stages. In the first stage, they generate 50 primary studies, each having an equal number of observations (either 25, 50, 100, 200, or 400 observations). Each primary study is described by the following DGP, where we try to maintain S&D's notation to facilitate comparison with their paper:

(1) $Y_i = 1 + x_{1i} + x_{2i} + \varepsilon_i,$

where $i = 1,2, \dots, N$, $N \in \{25,50,100,200,400\}$; $x_{1i} \sim N\left(0, \frac{1}{par^2}\right)$, $par \in \{1,3,9\}$; $x_{2i} \sim N(0,1)$; and $\varepsilon_i \sim N(0,1)$. A dataset is generated and OLS is used to estimate $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$, where β_1 is the effect of interest. The corresponding *t* value is converted to a PCC using Equation (2):

(2) $r_p = \dfrac{t}{\sqrt{t^2 + df}}$,

where $df = N - 3$. Note that par is a parameter that allows one to control the size of the *t*-statistic associated with $\hat{\beta}_1$, and hence the size of r_p. The three values of par(=1,3,9) correspond to a population mean value of r_p, ρ, equal to 0.7071, 0.3162, and 0.1104. This process is repeated until a meta-analysis sample of 50 PCC values are collected.

In the second stage, a DerSimonian and Laird random effects estimator (RE) is used to estimate ρ. Two different weights are used. One is based on the "correct" variance of the estimate of r_p as recently demonstrated by Aert & Goos[2].

(3) $S_1^2 = \dfrac{\left(1 - r_p^2\right)^2}{df}.$

The other is based on the "suboptimal" estimator that is commonly employed in the economics meta-analysis literature[1,3]:

(4) $$S_2^2 = \frac{(1-r_p^2)}{df}.$$

S_2^2 differs from the "correct" variance in that its numerator is the square root of the numerator in S_1^2. It can be shown that S_2^2 is the sampling variance of r_p when $\rho = 0$.[4,5]

For each experiment, two RE estimates of ρ are produced, one based on S_1^2 and one based on S_2^2. S&D follow the two-stage process above and simulate 10,000 meta-analyses for each of 15 experiments corresponding to the different combinations of $\rho \in \{0.7071, 0.3162, 0.1104\}$ and $N \in \{25,50,100,200,400)$.

III. Replication of S&D

We first demonstrate that we are able to replicate S&D's simulation results. TABLE 1 reproduces their Table 1. The first two columns identify the research design for the respective experiments. The remainder of the table is divided into three sections, reporting results for Bias, RMSE, and Coverage.

S&D's Table 1 provides the empirical support for their claim that the "suboptimal" estimator for the variance of PCC, S_2^2, produces superior results compared to the "correct" estimator, S_1^2. Across the board, using the suboptimal weights based on S_2^2 results in lower Bias, smaller RMSE, and coverage rates closer to 95%.

For example, when the true value of PCC, ρ, equals 0.7071 and all the primary studies in the meta-analysis sample have 25 observations, the RE estimator based on S_2^2 produces estimates that have an average bias of 0.0235, an average RMSE of 0.0279, and an average coverage rate of 84.90%. In contrast, the "correct" estimator, S_1^2, produces estimates that have an average bias of 0.0455, an average RMSE of 0.0479, and an average coverage rate of 14.27%. The statistical dominance of S_2^2 is true for every experiment in the table. Our efforts to replicate S&D are placed side-by-side to S&D's results and are identical except for minor Monte Carlo error.

IV. S&D's Research Design Is Not Well Suited for Their Experiment

The observant reader might have noticed in the previous description of S&D's research design that not only was the population value of β_1 homogeneous across primary studies, but the error terms all had the same variance. Since all primary studies included in a given meta-analysis also have an equal number of observations, they will share the same population values of $s.e(\hat{\beta}_1)$. In that case, when sampling variances are unknown and must be estimated, neither RE, FE, or UWLS is efficient. The optimal estimator is OLS. We demonstrate this empirically in TABLE 2.

As before, Columns (1) and (2) report details about the respective experiments. The first two columns of each section reproduce the Bias, RMSE, and Coverage values from TABLE 1. But there is now a third column to the right of those columns reporting OLS estimates (see Columns 5, 8, and 11). As can be clearly seen, for each of the 15 experiments, OLS dominates the two RE estimators on all three dimensions.

We are now in a position to explain S&D's results. Having noted that the optimal estimator is not RE but OLS, the reason the RE estimator based on S_2^2 performs better is because it is closest to OLS. This is evident in the last two columns of TABLE 2. These two columns report the coefficient of variation (CV) for S_1 and S_2, where CV = (standard deviation / mean) x 100%. The CV for S_2 is approximately half that of S_1. In other words, using the suboptimal estimate of the standard error produces weights that are more uniform than those using the correct estimate, and thus closer to the equal weights employed by OLS.

V. A Fairer Test

As shown above, S&D's simulations do not provide a fair test of the consequences of using the suboptimal estimator for PCC standard errors because their Monte Carlo data environments assume homoskedasticity and effect homogeneity in the primary studies. In this section we present results from additional simulations that build on S&D's research design but add heteroskedasticity and effect heterogeneity.

<u>Case 1: Homoskedasticity and Effect Homogeneity in the Primary Studies</u>. Case 1 is identical to the S&D simulations above. We focus on the case where all primary studies have 200 observations.

<u>Case 2: Heteroskedasticity and Effect Homogeneity in the Primary Studies</u>. Case 2 introduces heteroskedasticity by allowing primary studies to have differing numbers of observations, $N \in \{25,50,100,200,400\}$. S&D restricted all primary studies for a given meta-analysis to have the same sample size. We mix primary studies with different sample sizes in the same meta-analysis. 50 primary studies are included in each meta-analysis and they consist of equal numbers of studies with 25, 50, 100, 200, and 400 observations. Other than the mix in sample sizes, the DGP for the primary studies is the same as before:

(5) $$Y_i = 1 + x_{1i} + x_{2i} + \varepsilon_i,$$

where $i = 1,2,\ldots,N$; $N = 25,50,100,200,400$; $x_{1i} \sim N\left(0,\frac{1}{par^2}\right), par \in \{1,3,9\}$; $x_{2i} \sim N(0,1)$, and $\varepsilon_i \sim N(0,1)$. Note that the different sample sizes induce heteroskedasticity in the primary studies' estimated effects, $\hat{\beta}_1$.

<u>Case 3: Heteroskedasticity and Heterogeneity in the Primary Studies</u>. Case 3 adds effect heterogeneity directly into the DGP of the primary studies.

(6) $$Y_i = 1 + \beta_1 x_{1i} + x_{2i} + \varepsilon_i,$$

where $i = 1,2,\ldots,N$; $N = 25,50,100,200,400$; $x_{1i} \sim N\left(0,\frac{1}{par^2}\right), par \in \{1,3,9\}$; $x_{2i} \sim N(0,1)$, $\varepsilon_i \sim N(0,1)$, and $\beta_1 \sim N(1,1)$.

<u>The estimators</u>. Our analysis begins by comparing three base meta-analytic estimators for ρ: OLS, UWLS, and a third weighted least squares estimator based on random effects. For each of the latter two estimators, we compare versions with weights based on S_1^2 and S_2^2. We

include OLS because it is optimal given homoskedasticity and effect homogeneity (Case 1). UWLS is a conventional weighted least squares estimator that uses inverse variance weights, either $\frac{1}{S_1^2}$ or $\frac{1}{S_2^2}$. We include it because Stanley and others advocate for its use.[6,7] We also include a random effects version of UWLS because Case 3 introduces effect heterogeneity. It uses inverse variance weights $\frac{1}{(S_1^2+\tau^2)}$ or $\frac{1}{(S_2^2+\tau^2)}$. As Van Aert & Jackson show, this estimator is equivalent to the Hartung-Knapp method for random-effects meta-analysis and is a natural extension to S&D's UWLS estimator.[8] To distinguish the two WLS models, we refer to them as UWLS(FE) and UWLS(RE).

Before describing the experiments, we note one more difference with S&D. S&D calculate the population values of ρ by substituting population values of underlying parameters into Equation (2). This is problematic when r_p is characterized by heterogeneity. To address this problem, we generate a million observations of r_p for each experiment and use its mean to calculate Bias, RMSE, and Coverage.

The experiments. Following S&D, we calculate three population values for r_p corresponding to the scale parameter $par \in \{1,3,9)$. Given the three cases above, this yields 9 experiments. Each experiment generates 10,000 simulated meta-analyses, and each meta-analysis consists of 50 primary studies. We calculate Bias, RMSE, and Coverage for each of the five estimators: OLS, UWLS(FE-S_1^2), UWLS(FE-S_2^2), UWLS(RE-S_1^2) and UWLS(RE-S_2^2). As before, we report CVs for S_1 and S_2 to compare their variation.

The results. TABLE 3 reports the results for the experiments corresponding to Case 1. These results should be similar to S&D's since the DGP for each experiment is characterized homoskedasticity and effect homogeneity in the primary studies. Columns (1) to (3) report design details for the respective experiments. Columns (4) and (5) report the CV values for S_1 and S_2, and Columns (6) to (10) report the performance results for the respective estimators.

Consistent with TABLE 2, S_1^2 and S_2^2 display relatively little variation, with coefficients of variation ≤ 10%, and CV(S_2^2) approximately half that of CV(S_1^2).

In line with S&D's findings, UWLS(FE-S_2^2) is superior to UWLS(FE-S_1^2), and UWLS(RE-S_2^2) is superior to UWLS(RE-S_1^2) on the dimensions of Bias, RMSE, and Coverage. However, OLS is superior to them all.

For example, the average Bias for OLS is 0.0000 and the next closest is 0.0024 (cf. UWLS(RE-S_2^2)). The average RMSE for OLS is 0.0080, and the next closest is 0.0085 (cf. UWLS(RE-S_2^2)). And the average coverage rate for OLS is 95.08%, and the next closest is 92.72% (cf. UWLS(FE-S_2^2) and UWLS(RE-S_2^2)).

TABLE 4 reports results for Case 2, where primary studies are characterized by homogeneity in the effect size but heteroskedasticity in the estimated effects. The latter is reflected in the increased variation of S_1 and S_2, as seen in Columns (4) and (5). The respective CV values increase from an average of 5.4% and 2.7% in TABLE 3, to 52.1% and 50.5% in TABLE 4. The cause of this increase is that within any given meta-analysis there is now a mix of primary studies with observations having 25, 50, 100, 200, and 400 observations.

Turning now to the results in Columns (6) through (10), we see that OLS dominates on Bias and Coverage but not RMSE. For example, the average Bias for OLS is 0.0001. The next closest is 0.0046 for UWLS(FE-S_2^2) and UWLS(RE-S_2^2). The average coverage rate for OLS is 94.97%. The next closest is 89.25%, again for UWLS(RE-S_2^2). On the other hand, UWLS(FE-S_2^2) and UWLS(RE-S_2^2) dominate on RMSE, with the two estimators being virtually tied. This case provides an example where S_2^2 is superior to S_1^2 without OLS being more efficient. We shall have more to say about this case below.

TABLE 5 reports results for Case 3, which introduces both heteroskedasticity and heterogeneity into the primary studies. Heteroskedasticity is once again introduced through mixing primary studies of different sample sizes in the same meta-analysis. But now the population value of β_1 is constant within a study, but random across studies, $\beta_1 \sim N(1,1)$.

OLS again dominates. Average Bias for OLS is -0.0001. Next closest is UWLS(RE-S_2^2) with an average bias of 0.0063. Average RMSE for OLS is 0.0437. Next closest is UWLS(RE-S_2^2) with an average RMSE of 0.0447. And average coverage rate for OLS is 94.49%, with the next closest being UWLS(RE-S_2^2) with 93.16%. Given the larger degree of effect heterogeneity, we expect random effects to be superior to fixed effects, and that is what we see. But OLS is superior to all.

To summarize our results to this point, while we confirm S&D's conclusion that using weights based on a "suboptimal" estimate of the PCC standard error produces better results than those based on the "correct" estimate, in almost all cases the best approach is to employ no weights and instead perform meta-analysis using OLS. This result should not be too surprising, as others have also found that unweighted estimators can outperform weighted ones.[9,10]

VI. Bias-Variance Trade-Off Explains TABLES 4 and 5

It was easy to understand why S_2^2 did better than S_1^2 in the simulations underlying TABLES 1-3 where OLS is unbiased and efficient. What is not obvious is why it generally does better in the heteroskedastic and heterogeneous environments of TABLES 4 and 5.

S&D supply the explanation. As they note, "*all inverse-variance meta-analyses, whether S_1^2 or S_2^2, will 'positively' bias the meta-analysis estimator*". Inverse variance weights introduce bias because they disproportionately weight larger PCC values. This is clearly seen in Equations (3) an (4). As r_p becomes more positive, S_1^2 and S_2^2 decrease, causing $\frac{1}{S_1^2}$ and $\frac{1}{S_2^2}$ to increase so that larger values of r_p get more weight. In contrast, OLS does not have a bias problem. It is unbiased and consistent. Working against that, though, is the fact that heteroskedasticity favors estimators that give greater weight to observations that are more precisely estimated.

These two competing factors generate a bias-variance trade-off. In all of our experiments, the bias-variance trade-off favors estimators that use S_2^2 rather than S_1^2. While the former produces less precise estimates, they are also less biased. The previous experiments demonstrate that OLS can be a better choice than either. But there may be yet other estimators that make better trade-offs between bias and variance. The last two columns of TABLES 4 and 5 give one such example.

Define $\tilde{S}$ as follows:

$$(7) \qquad \tilde{S} = \frac{\left(1-\bar{r}_p^2\right)^2}{df}.$$

$\tilde{S}$ is called a "smooth estimator" because it uses averages.[11] The only difference between Equation (7) and Equations (3) and (4) is that r_p is replaced with its sample mean, $\bar{r}_p$. If we substitute $\tilde{S}$ in UWLS(FE) and UWLS(RE), we get FE and RE versions of UWLS(Smooth). These are reported in Columns (11) and (12) of TABLES 4 and 5.

While OLS performs best with respect to Bias, the two UWLS(Smooth) estimators also do well on this dimension. In addition, they have lowest average RMSE in TABLE 4 and admirable coverage rates. In TABLE 5, where the PCC values are characterized by effect heterogeneity, the RE version of the UWLS(Smooth) estimator dominates on RMSE, and performs well on both Bias and Coverage. The UWLS(Smooth) estimators, particularly the RE version, make a better trade-off between estimator bias and variance than the other estimators.

To summarize, the reason S&D's "suboptimal" variance estimator produces better outcomes in their simulations than the "correct" estimator is because it makes a better trade-off between bias and variance. However, it doesn't follow that it should be the meta-analyst's estimator of choice when working with PCCs. As demonstrated above, other estimators are available that may perform better.

VII. Testing for Publication Bias

While S&D recommend S_2 for meta-analyses using PCCs, they do not address its use as an explanatory variable in an Egger regression. To investigate this, we return to the DGP of S&D's Table 1 and repeat the experiments. For each simulated meta-analysis of 50 r_p values, we estimate a meta-regression that includes S_2 as an explanatory variable:

(8) $r_{ps} = \beta_0 + \beta_1 S_{2s} + \epsilon_s, s = 1, \dots, 50$; where

(9) $S_{2s}^2 = \frac{\left(1-r_{ps}\right)^2}{N-3}$, $N = 200$.

We estimate β_0 and β_1 using the following three estimators: OLS, UWLS(FE-S_2^2), and UWLS(RE-S_2^2). This process is repeated for a total of 10,000 simulated meta-analyses.

As there is no publication bias, all coefficients for S_2 should be close to zero and the estimated constant term should be approximately equal to the true value of ρ. TABLE 6 reports the results. The estimated values in the table are averages over 10,000 regressions. For example, when $\rho = 0.7071$ and OLS is used to estimate Equation (8), the average estimated values for β_0 and β_1 are 1.416 and -14.136. The corresponding average standard errors are 0.007 and 0.135. The average R^2 is 99.6%, and 100% of the estimated slope coefficients are statistically different from 0 at the 5% level.

Similar results obtain for the other experiments in the table. In every case, one would incorrectly conclude that the meta-analysis samples were characterized by substantial publication bias/small study effects. Given that there is no publication bias, it is clear that the estimated negative relationship between r_p and S_2 is entirely a mathematical consequence of the formula for S_2 in Equation (4).

Biased estimates of the slope coefficients produce biased estimates of the constant term. In all of the experiments of TABLE 6, the "publication bias adjusted" estimates of ρ, what is often called "effect beyond bias", are greater than 1, a nonsensical result. While the simulations in TABLE 6 assume $\rho > 0$, the same induced relationship between r_p and S_2 arises when $\rho < 0$ except that the sign of the bias reverses and becomes positive. If we repeat the experiments using S_1 rather than S_2, the results are marginally better, but still entirely unsatisfactory.

VIII. Conclusion

This study builds on the simulation framework of a recent paper by Stanley & Doucouliagos (2023). S&D use simulations to support a recommendation that meta-analyses using partial correlation coefficients (PCCs) should employ a suboptimal estimator of the PCC's standard error, denoted S_2, when constructing weights for fixed effect and random effects estimation. While we confirm their simulation findings, their simulations and recommendation may give meta-analysts a misleading impression.

S_2 performs better than the "correct" estimator because it does a better job of trading precision for bias. However, as demonstrated in this study, other estimators, including OLS, do an even better job. Thus S&D's findings should not be interpreted as an endorsement of the use of S_2 in meta-analyses with PCCs. This is underscored when one considers the use of S_2 in Egger regressions and FAT-PET analyses. As is well-known, and as we clearly demonstrate, the mathematical relationship between PCC and its standard error induces a bias that can grossly misrepresent the true extent of publication bias.[12,13]

To summarize our results, if there is no publication bias, there exist better meta-analytic estimators than those that base their weights on S_2. If there is publication bias, there is little evidence at the present time to guide recommendations for best practice. This is a topic that would benefit from further research, including performance comparisons of PCCs with Fisher's z.[14]

References

1. Stanley TD, Doucouliagos H. Correct standard errors can bias meta-analysis. *Research Synthesis Methods.* 2023, in press.

2. Aert VR, Goos C. A critical reflection on computing the sampling variance of the partial correlation coefficient. *Research Synthesis Methods.* 2023, in press.

3. Stanley TD, Doucouliagos H. Meta-regression analysis in economics and business. London: Routledge, 2012.

4. Olkin I, Siotani M. Asymptotic distribution of functions of a correlation matrix. In: Ikeda S., ed. *Essays in Probability and Statistics*. Tokyo: Shinko Tsusho. 1976 (pp. 235–251).

5. Aloe AM, Thompson CG. The synthesis of partial effect sizes. *Journal of the Society for Social Work and Research,* 2013. 4(4):390–405.

6. Stanley TD, Ioannidis JP, Maier M, Doucouliagos H, Otte WM & Bartoš, F. Unrestricted weighted least squares represent medical research better than random effects in 67,308 Cochrane meta-analyses. *Journal of Clinical Epidemiology.* 2023, in press.

7. Stanley TD, Doucouliagos H. Neither fixed nor random: weighted least squares meta-regression. *Research Synthesis Methods*, 2017. 8(1), 19-42.

8. Aert VR, Jackson D. A new justification of the Hartung-Knapp method for random-effects meta-analysis based on weighted least squares regression. *Research Synthesis Methods*, 2019. 10(4), 515-527.

9. Fletcher D, Dixon PM. Modelling data from different sites, times or studies: weighted vs. unweighted regression. *Methods in Ecology and Evolution*, 2012. 3(1), 168-176.

10. Shuster JJ. Empirical vs natural weighting in random effects meta-analysis. *Statistics in Medicine*, 2010. 29(12), 1259-1265.

11. Hunter JE, Schmidt FL. Estimation of sampling error variance in the meta-analysis of correlations: Use of average correlation in the homogeneous case. *Journal of Applied Psychology*, 1994. 79(2), 171-177.

12. Hamman EA, Pappalardo P, Bence JR, Peacor SD, Osenberg CW. Bias in meta-analyses using Hedges' d. *Ecosphere*, 2018. 9(9), e02419.

13. Nakagawa S, Lagisz M, Jennions MD, Koricheva J, Noble DW, Parker, TH., ... & O'Dea RE. Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods in Ecology and Evolution*, 2022. 13(1), 4-21.

14. Fisher RA. , R. A. 1921. On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron*, 1921. 1:1–32.

TABLE 1
Replication of S&D

Research Design		Bias				RMSE				Coverage			
		S&D		Replication		S&D		Replication		S&D		Replication	
ρ	n	S_1^2	S_2^2	S_1^2	S_2^2	S_1^2	S_2^2	S_1^2	S_2^2	S_1^2	S_2^2	S_1^2	S_2^2
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
0.7071	25	0.0455	0.0235	0.0455	0.0234	0.0479	0.0279	0.0479	0.0280	0.1427	0.8490	0.1429	0.8475
0.7071	50	0.0226	0.0112	0.0224	0.0111	0.0247	0.0151	0.0246	0.0151	0.4012	0.9478	0.4073	0.9470
0.7071	100	0.0111	0.0053	0.0110	0.0052	0.0132	0.0090	0.0131	0.0089	0.6601	0.9740	0.6634	0.9771
0.7071	200	0.0055	0.0025	0.0056	0.0027	0.0074	0.0057	0.0075	0.0057	0.8132	0.9870	0.8018	0.9878
0.7071	400	0.0028	0.0013	0.0027	0.0013	0.0045	0.0038	0.0045	0.0038	0.8862	0.9907	0.8826	0.9909
0.3162	25	0.0347	0.0172	0.0346	0.0173	0.0461	0.0336	0.0459	0.0335	0.7350	0.8977	0.7387	0.9031
0.3162	50	0.0179	0.0083	0.0177	0.0081	0.0265	0.0208	0.0264	0.0207	0.8299	0.9369	0.8372	0.9371
0.3162	100	0.0089	0.0040	0.0090	0.0041	0.0159	0.0137	0.0159	0.0137	0.8971	0.9516	0.8986	0.9534
0.3162	200	0.0045	0.0021	0.0045	0.0021	0.0103	0.0094	0.0102	0.0094	0.9270	0.9560	0.9257	0.9581
0.3162	400	0.0023	0.0011	0.0022	0.0010	0.0068	0.0065	0.0068	0.0064	0.9423	0.9612	0.9422	0.9624
0.1104	25	0.0127	0.0059	0.0127	0.0059	0.0361	0.0322	0.0355	0.0316	0.9100	0.9381	0.9135	0.9402
0.1104	50	0.0069	0.0030	0.0073	0.0034	0.0227	0.0211	0.0227	0.0211	0.9330	0.9517	0.9271	0.9473
0.1104	100	0.0035	0.0015	0.0036	0.0016	0.0151	0.0145	0.0149	0.0143	0.9429	0.9531	0.9447	0.9550
0.1104	200	0.0017	0.0007	0.0017	0.0007	0.0103	0.0101	0.0103	0.0101	0.9496	0.9552	0.9463	0.9511
0.1104	400	0.0008	0.0004	0.0010	0.0005	0.0071	0.0071	0.0070	0.0070	0.9509	0.9560	0.9536	0.9562
	Average	**0.0121**	**0.0059**	**0.0121**	**0.0059**	**0.0196**	**0.0153**	**0.0195**	**0.0153**	**0.7947**	**0.9471**	**0.7950**	**0.9476**

NOTE: Columns (1), (2), (3), (4), (7), (8), (11), and (12) are reproduced from Table 1 in Stanley & Doucouliagos (2023). They come from random effects estimates of ρ using PCC standard errors S_1 and S_2. Columns (5), (6), (9), (10), (13), and (14) are replications of S&D's results based on the code they provided with their paper. The table demonstrates that S_2 is superior to S_1.

TABLE 2
S&D's Results Explained

Research Design		Bias			RMSE			Coverage			CV	
ρ	n	S_1^2	S_2^2	OLS	S_1^2	S_2^2	OLS	S_1^2	S_2^2	OLS	S_1	S_2
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
0.7071	25	0.0454	0.0236	-0.0081	0.0479	0.0280	0.0178	0.1447	0.8441	0.9361	29.0%	14.8%
0.7071	50	0.0224	0.0111	-0.0037	0.0246	0.0151	0.0111	0.4082	0.9463	0.9430	20.1%	10.1%
0.7071	100	0.0111	0.0053	-0.0018	0.0131	0.0089	0.0074	0.6609	0.9772	0.9488	14.2%	7.1%
0.7071	200	0.0055	0.0026	-0.0009	0.0075	0.0057	0.0052	0.8124	0.9852	0.9481	10.0%	5.0%
0.7071	400	0.0028	0.0014	-0.0004	0.0045	0.0038	0.0036	0.8828	0.9915	0.9514	7.1%	3.5%
0.3162	25	0.0345	0.0172	-0.0064	0.0459	0.0335	0.0276	0.7400	0.9029	0.9465	13.2%	6.9%
0.3162	50	0.0178	0.0082	-0.0030	0.0264	0.0207	0.0187	0.8366	0.9368	0.9508	9.1%	4.7%
0.3162	100	0.0089	0.0040	-0.0015	0.0159	0.0137	0.0130	0.8943	0.9542	0.9506	6.3%	3.2%
0.3162	200	0.0044	0.0020	-0.0007	0.0102	0.0094	0.0091	0.9280	0.9583	0.9498	4.5%	2.3%
0.3162	400	0.0023	0.0010	-0.0003	0.0068	0.0064	0.0063	0.9441	0.9607	0.9516	3.2%	1.6%
0.1104	25	0.0128	0.0060	-0.0025	0.0355	0.0316	0.0288	0.9158	0.9435	0.9517	7.1%	3.7%
0.1104	50	0.0067	0.0029	-0.0013	0.0227	0.0211	0.0202	0.9323	0.9492	0.9523	4.1%	2.1%
0.1104	100	0.0035	0.0015	-0.0005	0.0149	0.0143	0.0140	0.9462	0.9560	0.9527	2.5%	1.3%
0.1104	200	0.0018	0.0009	-0.0001	0.0103	0.0101	0.0099	0.9513	0.9577	0.9527	1.7%	0.9%
0.1104	400	0.0008	0.0003	-0.0002	0.0070	0.0070	0.0069	0.9569	0.9603	0.9529	1.1%	0.6%
	Average	**0.0120**	**0.0059**	**-0.0021**	**0.0195**	**0.0153**	**0.0133**	**0.7970**	**0.9483**	**0.9493**	**8.9%**	**4.5%**

NOTE: Columns (1), (2), (3), (4), (6), (7), (9), (10) are reproduced from TABLE 1. The OLS results in Columns (5), (8), and (11) use the same meta-analysis datasets but estimate ρ from an OLS regression of r_p on a constant term. Columns (12) and (13) report the coefficient of variation (CV) of S_1 and S_2. The table demonstrates that unweighted, OLS regression is superior to weighted, random effects estimates and that the S_2 estimator is better than the S_1 estimator because it is closest to OLS.

TABLE 3
Comparison of Estimators Given Homoskedasticity and Homogeneity in the Primary Studies (Case 1)

	Design		CV		OLS	UWLS(FE)		UWLS(RE)	
ρ	n	τ^2	S_1	S_2		S_1^2	S_2^2	S_1^2	S_2^2
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					BIAS				
0.7062	200	0	10.0%	5.0%	0.0000	0.0071	0.0036	0.0064	0.0036
0.3155	200	0	4.5%	2.3%	-0.0001	0.0056	0.0027	0.0050	0.0026
0.1102	200	0	1.7%	0.9%	0.0001	0.0023	0.0011	0.0020	0.0011
		Average	**5.4%**	**2.7%**	**0.0000**	**0.0050**	**0.0025**	**0.0045**	**0.0024**
					RMSE				
0.7062	200	0	10.0%	5.0%	0.0050	0.0087	0.0061	0.0081	0.0061
0.3155	200	0	4.5%	2.3%	0.0091	0.0108	0.0095	0.0105	0.0095
0.1102	200	0	1.7%	0.9%	0.0099	0.0103	0.0100	0.0102	0.0100
		Average	**5.4%**	**2.7%**	**0.0080**	**0.0099**	**0.0086**	**0.0096**	**0.0085**
					COVERAGE				
0.7062	200	0	10.0%	5.0%	0.9510	0.7127	0.8928	0.7515	0.8928
0.3155	200	0	4.5%	2.3%	0.9485	0.9065	0.9377	0.9139	0.9378
0.1102	200	0	1.7%	0.9%	0.9530	0.9460	0.9511	0.9466	0.9510
		Average	**5.4%**	**2.7%**	**0.9508**	**0.8551**	**0.9272**	**0.8707**	**0.9272**

NOTE: The DGP underlying the simulations in this table are identical to the $\rho = \{0.7071, 0.3162.0.1104\}$, $N = 200$ experiments in TABLES 1 and 2. The main differences are (i) that ρ is calculated by simulating 1,000,000 values of r_p and taking their average, and (ii) it compares performance for the following estimators: OLS, UWLS(FE-S_1^2), UWLS(FE-S_2^2), UWLS(RE-S_1^2) and UWLS(RE-S_2^2). This table confirms the finding that OLS dominates the weighted estimators when primary studies are characterized by homoskedasticity and effect homogeneity.

TABLE 4
Comparison of Estimators Given Heteroskedasticity and Homogeneity in the Primary Studies (Case 2)

Design			CV		OLS	UWLS(FE)		UWLS(RE)		UWLS(Smooth)	
ρ	n	τ^2	S_1	S_2		S_1^2	S_2^2	S_1^2	S_2^2	FE	RE
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
BIAS											
0.7042	25,50,100,200,400	0	56.7%	51.7%	0.0000	0.0115	0.0064	0.0113	0.0064	0.0018	0.0018
0.3138	25,50,100,200,400	0	50.5%	50.0%	0.0000	0.0093	0.0052	0.0092	0.0051	0.0014	0.0014
0.1096	25,50,100,200,400	0	49.2%	49.7%	0.0003	0.0039	0.0023	0.0039	0.0023	0.0009	0.0009
		Average	**52.1%**	**50.5%**	**0.0001**	**0.0082**	**0.0046**	**0.0081**	**0.0046**	**0.0014**	**0.0014**
RMSE											
0.7042	25,50,100,200,400	0	56.7%	51.7%	0.0095	0.0130	0.0086	0.0128	0.0086	0.0061	0.0061
0.3138	25,50,100,200,400	0	50.5%	50.0%	0.0164	0.0142	0.0116	0.0143	0.0117	0.0104	0.0106
0.1096	25,50,100,200,400	0	49.2%	49.7%	0.0178	0.0123	0.0116	0.0125	0.0117	0.0113	0.0114
		Average	**52.1%**	**50.5%**	**0.0146**	**0.0131**	**0.0106**	**0.0132**	**0.0107**	**0.0093**	**0.0094**
COVERAGE											
0.7042	25,50,100,200,400	0	56.7%	51.7%	0.9458	0.5197	0.8016	0.5670	0.8016	0.9376	0.9418
0.3138	25,50,100,200,400	0	50.5%	50.0%	0.9509	0.8664	0.9222	0.8825	0.9255	0.9493	0.9544
0.1096	25,50,100,200,400	0	49.2%	49.7%	0.9524	0.9400	0.9475	0.9455	0.9504	0.9514	0.9549
		Average	**52.1%**	**50.5%**	**0.9497**	**0.7754**	**0.8904**	**0.7983**	**0.8925**	**0.9461**	**0.9504**

NOTE: The simulations underlying this table are identical to those in TABLE 3 except that the DGP of the primary studies are characterized by heteroskedasticity and effect homogeneity as described by Equation (5) in the text. In addition to the estimators in TABLE 3, the table also reports estimates from two additional estimators: an unrestricted WLS Fixed Effect estimator, and an unrestricted WLS Random Effects estimator, where the respective PCC standard errors are given by the "smooth estimator" of Equation (7). While OLS dominates on Bias and Coverage, the weighted regressions using S_2 are superior to OLS on RMSE. However, the two UWLS(*Smooth*) estimators have lowest RMSE.

TABLE 5
Comparison of Estimators Given Heteroskedasticity and Heterogeneity in the Primary Studies (Case 3)

	Design		CV		OLS	UWLS(FE)		UWLS(RE)		UWLS(Smooth)	
ρ	**n**	**τ^2**	**S_1**	**S_2**		**S_1^2**	**S_2^2**	**S_1^2**	**S_2^2**	**FE**	**RE**
(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
					BIAS						
0.5115	25,50,100,200,400	1	80.6%	61.1%	0.0001	0.3273	0.1950	0.0233	0.0105	0.0005	0.0017
0.2810	25,50,100,200,400	1	53.9%	51.1%	-0.0001	0.0982	0.0452	0.0095	0.0062	0.0008	0.0005
0.1076	25,50,100,200,400	1	49.4%	49.7%	-0.0003	0.0085	0.0042	0.0043	0.0022	0.0004	0.0001
		Average	**61.3%**	**54.0%**	**-0.0001**	**0.1447**	**0.0815**	**0.0124**	**0.0063**	**0.0006**	**0.0008**
					RMSE						
0.5115	25,50,100,200,400	1	80.6%	61.1%	0.0671	0.3345	0.2103	0.0790	0.0708	0.0886	0.0677
0.2810	25,50,100,200,400	1	53.9%	51.1%	0.0406	0.1170	0.0721	0.0431	0.0418	0.0515	0.0404
0.1076	25,50,100,200,400	1	49.4%	49.7%	0.0233	0.0263	0.0245	0.0223	0.0215	0.0235	0.0210
		Average	**61.3%**	**54.0%**	**0.0437**	**0.1593**	**0.1023**	**0.0481**	**0.0447**	**0.0545**	**0.0430**
					COVERAGE						
0.5115	25,50,100,200,400	1	80.6%	61.1%	0.9365	0.0275	0.1915	0.8716	0.9154	0.8447	0.9342
0.2810	25,50,100,200,400	1	53.9%	51.1%	0.9475	0.3779	0.6936	0.9271	0.9361	0.8528	0.9463
0.1076	25,50,100,200,400	1	49.4%	49.7%	0.9507	0.8571	0.8769	0.9403	0.9434	0.8830	0.9493
		Average	**61.3%**	**54.0%**	**0.9449**	**0.4208**	**0.5873**	**0.9130**	**0.9316**	**0.8602**	**0.9433**

NOTE: The simulations underlying this table are identical to those in TABLE 4 except that the DGP of the primary studies are characterized by heteroskedasticity and effect heterogeneity as described by Equation (6) in the text. The table demonstrates that OLS dominates the weighted estimators using S_1 and S_2. However, the UWLS(*Smooth-RE*) estimator has lowest RMSE.

TABLE 6
FAT/PET with S_2

Research Design ρ n	Estimated Equation	Type I Error Rate
	OLS-S_2	
0.7071 200	$r_p = 1.416 - 14.136\ S_2$, $R^2 = 0.996$ (0.007) (0.135)	100%
0.3162 200	$r_p = 3.124 - 41.743\ S_2$, $R^2 = 0.978$ (0.060) (0.891)	100%
0.1104 200	$r_p = 7.780 - 108.847\ S_2$, $R^2 = 0.841$ (0.479) (6.791)	100%
	UWLS(FE-S_2)	
0.7 200	$r_p = 1.409 - 14.011\ S_2$, $R^2 = 0.996$ (0.007) (0.133)	100%
0.3 200	$r_p = 3.100 - 41.390\ S_2$, $R^2 = 0.978$ (0.059) (0.878)	100%
0.1 200	$r_p = 7.731 - 108.159\ S_2$, $R^2 = 0.843$ (0.472) (6.705)	100%
	UWLS(RE-S_2)	
0.7 200	$r_p = 1.409 - 14.011\ S_2$, $R^2 = 0.996$ (0.007) (0.133)	100%
0.3 200	$r_p = 3.101 - 41.406\ S_2$, $R^2 = 0.978$ (0.059) (0.879)	100%
0.1 200	$r_p = 7.735 - 108.209\ S_2$, $R^2 = 0.843$ (0.473) (6.712)	100%

NOTE: The DGP underlying the simulations in this table are identical to the $\rho = \{0.7071, 0.3162, 0.1104\}$, $N = 200$ experiments in TABLE 2. However, instead of estimating ρ, the estimated r_p and S_2 values are used to estimate Equation (8). The values in the table are averages of estimated parameters from 10,000 meta-analyses. The table demonstrates the extent to which the mathematical relationship between r_p and S_2 given by Equation (4) causes Egger regressions/FAT-PET analyses to misestimate the existence of publication bias and the “effect beyond bias”.