Received: 21 April 2022 | Revised: 16 July 2022 | Accepted: 12 September 2022

DOI: 10.1111/kykl.12321

ORIGINAL ARTICLE

KYKLOS WILEY

Meta-mining: The political economy of meta-analysis

Martin Paldam

Department of Economics and Business, Aarhus University, Fuglesangs Allé 4, Aarhus V, 8210, Denmark

Correspondence
Martin Paldam, Department of Economics and Business, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark.
Email: mpaldam@econ.au.dk. Net: http://www.martin.paldam.dk/.

Abstract

Meta-analysis studies the literature reporting estimates of one parameter, which at present is assumed positive. The purpose of the analysis is to find the best meta-average, which corrects the mean of the estimates for bias. The two main biases are: (i) Publication bias, where the correction nearly always makes the average smaller. (ii) Omitted variable bias, where the correction typically makes the average larger. Consequently, the bias is likely to increase if the correction is for the wrong bias. This allows a game of meta-mining to be played. A case study demonstrates the scope for meta-mining, and that it has been done. The game of meta-mining is surely against the purpose of meta-analysis.

KEYWORDS
control variables, data mining, meta-analysis, replication

JEL CLASSIFICATION
B4, C12, O19

1 | INTRODUCTION

All sciences know that results need repeated replication to be believable. Mueller-Langer et al. (2019) find that only 0.1% of economics papers are replicated, so replication is rare, and when done, it frequently gives embarrassing results, as in section 4.4.[1] However, the same effect, β, is often studied in a β-literature of many papers. Instead of

[1]Google scholar has almost 150,000 hits to "replication crisis in economics" (March 2022).

strict replication, the profession faces a swarm of partial replications. The results of the primary studies often have a large range.[2]

Once the number of papers is large, such as 100, the β-literature represents a big effort such as 30 man-years, trying to 'catch' β in all the ways the authors have thought about. The data often overlap, and control variables go in and out in different combinations. Researchers often try many model variants to find the 'right' one, and many results remain unpublished. ***Meta-analysis*** is a technique to summarize this effort. Thus, it aims at replacing replication.

Researchers are humans with priors like everybody else. That is, we have both preferences and interests, and as research requires choices, it will be influenced by the priors. If they differ randomly, their effect on the average result is random too. However, common preferences or a large sponsor who affects the interests of many researchers in the same way may bias a whole literature. Meta-studies show that such biases are common – notably when β is the effect of a policy where political and bureaucratic interests enter, as discussed in a moment.

Meta-analysis starts from the ***mean*** of the results and tries to detect a ***bias***. It is often found, and it then estimates a ***meta-average*** corrected for the bias. Consequently, meta-analysis reduces the bias due to common priors of researchers. Approximately one thousand meta-studies have been made in economics; see e.g., Doucouliagos and Paldam (2006). Also, a good textbook, Stanley and Doucouliagos (2012), is available. For new readers of meta-analysis, the Appendix provides the basic intuition behind the FAT-PET MRA, which is the main meta-tool used below.

This paper takes the next step. It notes that meta-analysts have priors too, which may cause ***meta-bias***. The purpose of this paper is to develop the (new) theory of meta-bias. It demonstrates how meta-bias can be detected and avoided. A case study analyzes the range of meta-averages. It is less than the range of the primary studies, but it is still substantial.

1.1 | The political economy of estimating policy effects[3]

Economic studies often deal with the effect of a policy. It always has an announced goal shared by many, and some theory proposes that the policy may serve this goal. However, the theory is qualitative and needs a quantitative estimate, which is provided by the β-literature. The following argues that there are good reasons to expect that such literatures have an upward bias.

It is likely that researchers with the strongest preferences for the β-goal are most attracted to the field. The policy is implemented by a public bureau with a budget. The classical model of such bureaus (Niskanen, 1994) suggests that the bureau wants to grow, so it wants estimates of the policy effect to be large – too large is better than too small. The budget often sponsors research either directly or by rewarding loyal researchers in other ways. Obviously, the bureau prefers to sponsor researchers who are friendly towards its goal, and those sponsored by a bureau may develop loyalty. Consequently, collusion may result.

It is common that the main sponsor of research in such cases is the bureau. Thus, many researchers have both a preference for 'good' results and an interest in such results. As regressions are cheap to run once the data are in the machine, it is rational to mine the data by running many regressions and choose the best, which is likely to be too good.

The case study in section 4 is a typical case. It considers the first 141 studies of development aid effectiveness, where well-replicated meta-studies exist. Here both preferences and interests come together. We all prefer aid to work, so that poverty in the world is reduced. Aid budgets also sponsor research in development. In addition, most

[2]One reason that researchers should make meta-studies is that it is sobering to discover the width of the range of results even when the results are published in perfectly decent journals.

[3]A general theory of the research of the representative economist is found in Paldam (2018). At present, the analysis concentrates on the political economy.

researchers in development have interests in the large aid industry, which employs many consultants, notably in project missions that are well-paid and require economists.

1.2 | The basis for meta-mining: two biases in the opposite direction

The β-literature consists of all papers with estimates (b, s) of β and its standard error.[4] The (b, $1/s$)-scatter is the ***funnel***, which shows the distribution of the estimates of the β-literature. It has a broad base at low precision, and as precision $p = (1/s)$ rises, it narrows. When it is asymmetric, it points to bias. The meta-average is at the axes of symmetry once it is corrected for the asymmetry. The most precise estimates are likely to be close to this axis. Thus, the meta-average is found by weighting the estimates by their precision; see Appendix.

Results vary mainly because estimates include control variables in many combinations. Therefore, the perspective of the paper is one control variable, $\boldsymbol{\zeta}$ (zeta), which is included in some, but not all, estimates, as indicated by the binary (0, 1) inclusion variable ***z***. The effect of ζ on β is β_z, where the estimates are b and b_z respectively. The probability, π, that ζ is included in the estimating equation is taken as a function of b_z, i.e., $\pi = \pi\ (b_z)$.

The paper looks at the two main biases: ***Publication bias*** occurs if ζ is ***systematically*** included for its effect b_z on b, i.e., $\partial\pi/\partial b_z > 0$. If b_z is insignificant or 'wrong' (negative), the estimate has a high probability of being censored. Thus, the publication bias is positive, and the meta-average corrected for the bias is smaller than the mean. ***Omitted variable bias*** occurs if ζ is included independently of b_z, i.e., $\partial\pi/\partial b_z \approx 0$. Normally, the variable improves the results. Thus, the publication bias is negative. It follows that if a meta-study adjusts for the wrong bias, notably for a false omitted variable bias, the bias increases. This introduces the theory of ***meta-mining***, which is the main theme of this paper.

The paper proceeds as follows: Section 2 discusses why estimating equations contain control variables. Section 3 considers the two biases and how to adjust for them. For ease of presentation, section 2 and 3 assumes that β has ***one true value***[5] that ***is positive***.[6] Section 4 is a case study showing the range of results that can and have been reached by meta-mining. A Net Appendix brings extra tables to this section. Section 5 concludes that meta-mining defeats the purpose of meta-analysis. The Appendix is a brief introduction the FAT-PET tool of meta-analysis and presents a table that should help the reader keep track of the variables discussed.

2 | WHY ARE ADD-ON VARIABLES INCLUDED IN MODELS?

This section is fueled by two observations: (i) I have looked at many meta-studies. Meta-studies normally report a funnel diagram displaying the distribution of the estimates, i.e., the $b \approx \beta$. Published estimates have fine t-values, but funnels are still amazingly wide. As mentioned, most of this variation is due to control variables that go in and out of models. (ii) I have made many primary studies myself, and I know that estimates often react substantially to the inclusion/exclusion of controls.

Apart from variables coming from the theoretical model, most estimation models contain add-on variables. The two main types are ***ad hoc*** controls and ***ceteris paribus*** controls meant to control for sample differences. Section 2.1 discusses the concept of an *add-on control*, while section 2.2 looks at control variables that should be in some models

[4]Three points should be made: (i) The meta-analyst should not prejudge papers. A breakthrough claimed by one paper may not replicate in the next; see section 4.4. (ii) It has often been investigated whether the impact factor of the journal matter for the result – this is normally rejected. (iii) The typical paper reports about 10 estimates. The resulting clustering of results can be handled by several methods that are not discussed at present.

[5]Meta-studies show if the literature finds a robust and unbiased estimate of β. We like to believe that such meta-results are close to the truth – or at least that they are closer to truth than the mean.

[6]Everything generalizes to a negative effect. The case where the theory claims that there is no effect is not discussed.

but not in others. Section 2.3 discusses ad hoc controls, which seem to be the most common add-on variables. Section 2.4 asks if an ad hoc variable belongs.

2.1 | *Theory and operationalization with one add-on variable ζ ($=\zeta_1$)*

The theoretical model explains y by x and some other variables such as q_1 and q_2:

(1a) $y = F(x, q_1, q_2)$, which is linearized as (1b) $y = \alpha + \beta x + \lambda_1 q_1 + \lambda_2 q_2$

The estimated equation typically looks as follows:

(2a) $y = \{\alpha + \beta_1 x + \gamma \zeta\} + [] + u$, model with one control ζ and other controls in []
(2b) $y = \{\alpha + \beta_2 x\} + [] + u$ as (2a), without ζ. The effect of ζ is $b_z = b_1 - b_2$
(2c) $[] = [\gamma_2 \zeta_2 + ... + \gamma_n \zeta_k]$ other controls, which include q_1 and q_2

[] holds the remaining $k - 1$ control variables. The theory includes q_1 and q_2, so they should be included in all estimates. The remaining ζ's are add-on variables. The effect of ζ on b is b_z, while the coefficient to x in (2b) is b_1, which is $b_1 = b_2 + b_z$ in (2a). The residuals are u. Obviously, both b_1 and b_z are estimates and depend upon the data sample and the other controls [].

The add-on variable ζ is used in some of the papers, but not in all. Journals economize on space, and authors often exercise self-control or are told to reduce their papers. Thus, authors may omit to mention ζ if $b_z \approx 0$ in the sample used. Perhaps a footnote will say that the author also tried variable ζ, but that it did not work. Such notes may be overlooked by the meta-analyst.

2.2 | Two types of variables that should be in some studies but not in others

The variables that should not be in all studies are ceteris paribus controls and alternative representatives of the same latent variable.

Ceteris paribus controls: Each dataset used in the literature differs, as it has some general and some specific traits. To reach a general estimate, it should include variables that control for the specific traits. Such controls are *cp-controls* (for ceteris paribus). They do not give bias but reduce variation. The cp-controls should make the funnel leaner, but as mentioned, few funnels are lean. Hence, most controls cannot be cp-controls.

Example: Many cross-country estimates contain an Africa-dummy to tackle the African problems of bad climate, low levels of education, poor infrastructure, etc. It normally gets a negative coefficient, and consequently b_z is positive. It is obviously wrong to treat the Africa-dummy as an omitted variable in estimates covering other countries. It might be omitted in some papers that include African countries, but then these studies are likely to include explicit variables for the African problems. The example of section 4 includes an Africa-dummy.

Alternative variables: Many variables are confluent. If two variables, ζ_1 and ζ_2, contain a common factor that affects b, it is likely that the papers include either ζ_1 or ζ_2.[7] If ζ_1 is taken as an omitted variable when it is missing, and ζ_2 is taken as an omitted variable when it is missing, the effect of the common factor is counted twice, and hence the estimate of b is biased. Only if the estimate contains neither ζ_1 nor ζ_2, is it missing a variable bias.

[7]The right technique is to extract the common factor and use it as the explanatory variable. When a choice between z_1 and z_2 is made, the choice may be too good.

Example: Cross-country studies normally consider many confluent variables because development is a process that causes transitions in most variables,[8] so confluence is a large problem. Thus, the adjustment for omitted variables may generate rather large biases; see Paldam (2021).

2.3 | The pool of ad hoc controls: reasonable, but not inevitable controls

Most control variables are not of the types mentioned but are ad hoc controls that are added for their effect after a search by the researcher. In most fields, the literature has generated a set of K 'permissible' controls. These are controls which may be added at the author's discretion. As they have already been used in other papers, they do not need a thorough discussion. Such controls seem to be the most common ones. Section 4 considers a literature with a pool of 22 such variables; this appears to be a typical number. A researcher may try all or some of these, and if the dataset has, e.g., $n = 250$ observations, the estimating model may have five controls. Five controls can be chosen from the 22 possible ones in $\binom{22}{5} = 26{,}334$ ways. Each choice gives an estimate of β. No researcher tries all these estimates, but it is easy to make, e.g., 1,000 estimates in a search process lasting a couple of days, and various diagnostic statistics increase the efficiency of the search.

If the dataset had $n = 1{,}000$ observations, all 22 ad hoc variables could be thrown in and tested down so the ones working could be determined. There is always some randomness in the process, but if n is large, the random element is small. If the process is done with $n = 100$, the randomness may be substantial.

In addition, a paper may introduce a new control. This may even justify the paper. Thus, when a new variable is introduced, it has a big effect, but later papers often show that the effect is exaggerated; see section 4.4 for two such cases.

I have lectured many articles for my students, and in so doing, I have often presented a slide with the control variables of the article and asked the students why these controls were chosen. This always gives some discussion. I believe that it is common that the choice is directed by the desire to make the result better in three ways: (i) It gets closer to the prior of the researcher; (ii) it gets closer to the theoretical predictions, which may be the same as (i); and (iii) it becomes more significant statistically.

2.4 | Does an ad hoc variable ζ belong in the estimation equation?

If b_z is significant when the sample size goes to infinity, ζ belongs in (2). If b_z goes to zero, the control does not belong. If β_z is a sizable effect, it is likely to be discovered early in the literature, and ζ will be included from then on, which is discussed as the learning case in section 3.3 below. If β_z is small, many of the b_zs will be insignificant and remain unreported.

Figure 1 shows how the stylized distribution of b_z should look in two cases. They look like the typical funnel graph shown in Figure 2 below, but they have different scales at both axes, and the horizontal axis is different. Figure 1a is the case where ζ does not belong in the model, but sometimes becomes positive by chance. If only significantly positive estimates are published, and it is assumed that this is a case of an omitted variable, so that the correction made assumes that the average effect of ζ is a, the meta-average becomes biased by a.

Figure 1b is the case where ζ should be in the model, but sometimes it becomes insignificant by chance. If β_z is substantial as shown, only a few estimates have insignificant and thus unpublished ζ-effects. The average of the significant coefficients is a. Thus, if all studies have tried ζ, but only the studies in which ζ becomes significant are published, the estimates of b_z are relatively small when they are insignificant. Thus, a is a biased estimate of $\bar{b}_z$.

Few authors write which controls are ad hoc controls and which are cp-controls.

[8]A transition is a systematic change in the variable, so that it diverges from the traditional level as poor countries start to develop and converge to a different level in the modern countries. Transitions typically give correlations of 0.5 to 0.8 of socio-economic variables to income, in wide cross-country data samples; see Paldam (2021).

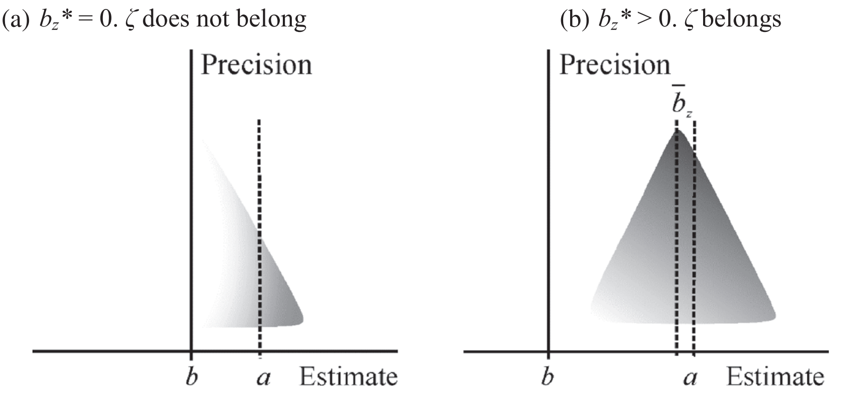


FIGURE 1 Funnel graphs for the effect b_z of ζ on the estimate of β. Funnel of b_z, where n is taken to be large. The gray shading indicates that the effect is significantly positive. Such estimates have a (much) larger chance of publication. The average of the significant b_z's is a.

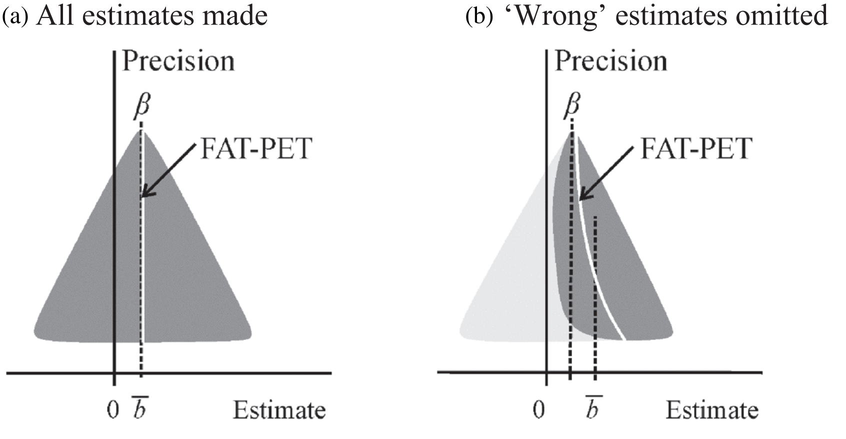


FIGURE 2 Stylized funnel graphs illustrating ***publication bias***. Figure 2a is symmetric, and $\bar{b} = \beta = b_M$, so the FAT-PET is vertical. Figure 2b occurs when researchers and journals refrain from publishing estimates that have the wrong sign or are insignificant, so that the light gray part of the funnel is suppressed. It biases the mean. As p increases, the FAT-PET converges to $b_M = \beta < \bar{b}$.

3 | PUBLICATION BIAS AND OMITTED VARIABLE BIAS

Sections 3.1 to 3.3 look at publication bias, omitted variable bias and the case of learning. Sections 3.4 and 3.5 argue that if a publication bias is detected at the basic level by a FAT-PET (see Appendix), it is dangerous to augment the FAT-PET, as it increases the PET and decreases the FAT. Thus, the bias returns, and the tool used to detect the bias is blunted.

The funnel plot has b at the horizontal axis and p ($= 1/s$) at the vertical axis. The formula $b = b_M + b_F/p$, for the FAT-PET, explains b with precision p. It is a curve in the funnel plot. The coefficient on $1/p$ is typically *positive*, and hence the slope on the FAT-PET is negative. Thus, we expect to see the FAT-PET curve as a hyperbola, where b converges to the meta-average b_M as p rises, i.e., the FAT-PET converges to the axis of symmetry in the uncensored funnel.[9]

[9]Publication bias gives asymmetry due to censoring of unwanted results. The FAT-PET is meant to converge to the symmetry axis of the uncensored funnel. The top of the funnel should be close to the axis. Many simulations in which the true value is known show that the FAT-PET does what it is supposed to rather neatly.

3.1 | Publication bias

Publication bias means that the published results differ *systematically* from the true value. If the typical researcher behaves as predicted by economic theory, he will run many regressions and choose the best as discussed. This gives a publication bias, which can be detected and corrected by the FAT-PET.

Control ζ contributes to the bias if it is included as a *function* of b_z. Most estimates where b_z is negative or insignificant remain unpublished. Thus, published estimates are too large. Publication bias is corrected by giving high precision estimates of b a higher weight. This normally gives a *downward* correction of the mean, as illustrated in Figure 2b, where β is about half the mean, $\bar{b}$. It is a good rule-of-thumb to expect that the mean estimate $\bar{b} \approx 2\beta$; see Ioannidis et al. (2017) and Doucouliagos et al. (2018).

If the funnel has asymmetries, which can be explained as the result of censoring of weak or wrong estimates, it suggests publication bias. The FAT-test statistic $b_F > 0$, so the FAT-PET curve is negatively sloped, as shown by Figure 2b. This confirms the suggestion.

3.2 | Omitted variable bias

Control ζ generates an omitted variable bias if ζ is *randomly* included relative to b_z, and b_z is significant in the typical study. If b_z is substantial, this may give a funnel two tops, which differ by the average of b_z; see Figure 3. Assuming that b_z is positive, also at the limit β_z, so that ζ should be included, then β_2 is the right estimate and β_1 is biased, and so is the mean $\bar{b}$ as shown. This is the omitted variable bias, and it is (normally) negative. An omitted variable is corrected by giving estimates including ζ a larger weight. This amounts to adding $\bar{b}_z$ to the estimates not including ζ, and thus it causes an *upward* correction of the mean, which ideally moves the meta-average b_M from $\bar{b}$ to β_2.

In Figure 3a, it is a toss-up whether this is β_1 or β_2. If ζ should be included and $b_z > 0$, it means not only that b becomes larger, but also that the standard error of the estimate decreases, so that precision rises. Thus, it is likely that the β_2-peak is higher than the β_1-peak, and therefore the correction for publication bias will find that peak. This is especially likely if the estimates have publication bias in addition to the omitted variable bias, as shown in Figure 3b.

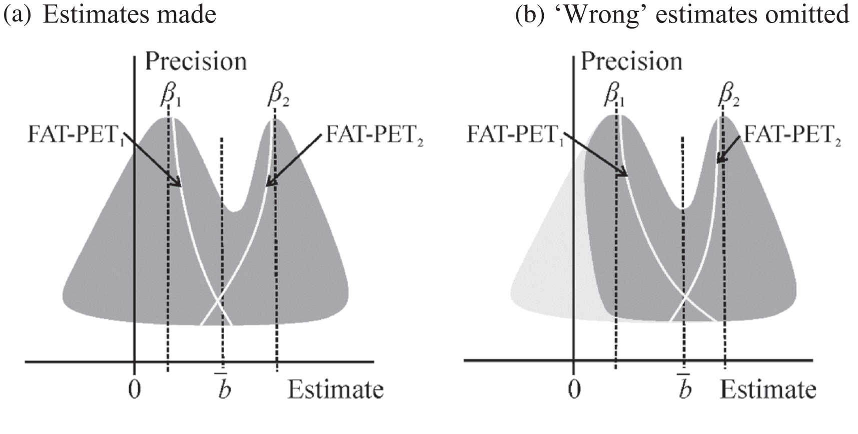


FIGURE 3 Stylized funnel graphs illustrating ***omitted variable bias***. The effect of ζ is $\beta_z = \beta_2 - \beta_1$. If ζ should be included in all estimates, β_2 is the true value and thus $\bar{b} < \beta$, both on Figure 3b and especially on Figure 3a. The graph assumes that ζ has the probability of about 0.5 for being included.

The technique to handle an omitted variable, ζ, is to augment the FAT-PET with the z-inclusion variable, where the meta-average is termed b_A. If ζ is randomly included, b_A gives an unbiased estimate of β, and in addition, it finds the average value of b_z. If the bias is negative, the augmentation with z increases the estimate.

3.3 | Omitted variable with learning

Imagine that ζ was not included from the start of the β-literature but was found to work in paper j. Researchers of new papers should read the old papers, so they would know that control ζ works after paper j, and a bit later it will be included in all papers, as illustrated by Figure 4a. In this case, the peak points to the true value, and corrections for either omitted variable bias or publication bias will find the same value, $b_M \approx \beta_2 > \bar{b}$.

If learning is combined with publication bias, as shown in Figure 4b, it is likely that the correction for publication bias – shown by the FAT-PET curve – is better at finding β_2 than a correction for omitted variable bias. Note that the FAT-PET has a positive slope, while its slope was negative in Figure 2b. However, it is likely that the mean, $\bar{b}$, is close to β_2 anyhow.

In the case of learning, imagine that the literature finds that ζ affects the result downward, so that the funnel shifts to the left. This will largely resemble Figure 3b, and the correction for publication bias will get close to the true value. Here the FAT-PET may have an insignificant slope. If the FAT-PET has a negative slope, it is a (strong) indication of a classical publication bias, and if the slope is positive, it points to a learning process. However, negative slopes are much more common than positive slopes in meta-studies.

3.4 | What is known if the basic level shows a publication bias?

About two thirds of all meta-studies (in economics) find a funnel asymmetry that indicates a publication bias. If $\beta > 0$, the FAT-PET will have a negative slope. Thus, a fraction of the published estimates has controls that are selected from their effect on b. The simple fact that estimates with the wrong sign (i.e., negative) are difficult to publish might be enough.

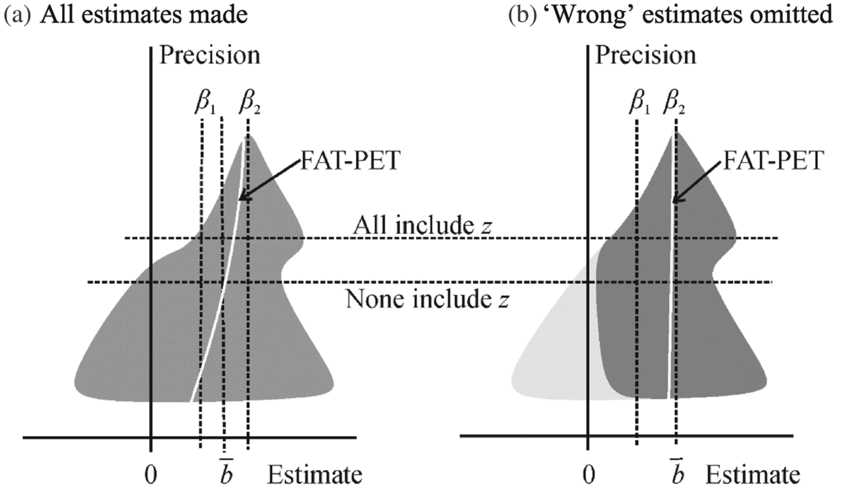


FIGURE 4 Stylized funnel graph for ***omitted variable with learning***. Figure 4a combines two funnels: the low is for the estimates without ζ. The peak is β_1. The upper funnel includes ζ. Here the funnel is shifted by $d_z = \beta_2 - \beta_1$. Between the two horizontal lines, some papers include ζ, and others do not. Figure 4b shows the situation from Figure 4a when wrong or insignificant estimates are censored. Here it is unclear where the mean, $\bar{b}$, is relative to β_2, and the FAT-PET is likely to be steep, as a is close to zero.

The β-literature has tried K controls, and we know that most of these controls are systematically selected for the size of b_z. If one such control is treated as an omitted variable, and the estimate is adjusted accordingly, the bias increases. This is the basis for meta-mining. It is surely wrong. The case study in section 4 considers a literature with a substantial publication bias. It has $K = 22$, of which one third are found in the average estimate, in many combinations.

In such cases it is, of course, possible that a few of the ζs are randomly included, but it is difficult to detect these variables. The random part of the effect of ζ will cause ζ not to work – and hence be excluded – randomly. This will give publication bias, not omitted variable bias.

3.5 | What happens to the FAT and the PET when the MRA is augmented?

Imagine you go ahead augmenting the FAT-PET when a publication bias has been found, i.e., $\beta_M < \bar{b}$. Assume further that ζ is one of the variables generating the bias, so that ζ is included when b_z is positive and significant only. This is coded by the inclusion variable z that is 1 when ζ is included and otherwise 0.

When $b = b_M + b_F s$ is augmented to $b = b_A + b_F s + gz$, both the PET and the FAT are estimated as if ζ was included in all estimates with its average effect when included. Thus, the PET increases, and the part of the publication bias caused by ζ disappears. Consequently, the FAT decreases. This is double bad. Not only does the publication bias increase, the tool used to detect the bias is also blunted.

By an augmentation with a handful of controls, one may even increase the PET to exceed the mean, so that $\beta_M < \bar{b} < \beta_A$. In addition, the FAT may become insignificant. In the case study in section 4, these effects prove to be substantial and hence easy to misuse.

4 | A CASE STUDY OF META-MINING

The case study uses the data for a meta-study of 141 papers with 1,779 estimates of the effect of development aid on growth; see Doucouliagos and Paldam (2006; 2008; 2011; 2013; 2015). The vast effort to find this effect is probably caused by the fact that researchers know that there is a problem: The simple correlation between aid and growth is zero; see Paldam (2022b).

Section 4.1 gives the basic meta-analysis, showing a publication bias as predicted by section 1.1. Section 4.2 shows how the PET and FAT react to augmentations (*aug*). Section 4.3 reports the scope for meta-mining. Section 4.4 turns to the issue of replication by looking at the two most cited AEL papers in the 21st century. A Net-Appendix (Paldam, 2022a) reports further evidence.

4.1 | The AEL, aid effectiveness literature: The basic meta-analysis

Due to the said problem, researchers have made large efforts to find a positive result. Many methods are available to put structure on data to make them say something "more" than the correlation. The application of these methods has resulted in many nicely significant positive results, as seen from the funnel on Figure 5. The main tool used to go from zero correlation to the wild scatter of Figure 5 is to add ad hoc controls to the estimating equation.

Table 1 reports the basic meta-analysis, reducing the mean of 0.07 to the meta-average of 0.03. Thus, the AEL has a substantial publication bias, it is even a little larger than two. Thus, the FAT-PET should not be augmented. The following two sectors show what happens when augmentations are nevertheless done.

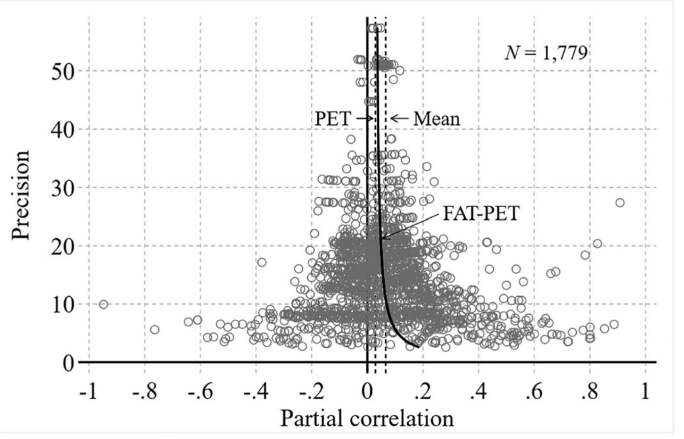


FIGURE 5 The funnel of the 1,779 estimates of aid effectiveness

TABLE 1 The mean and the basic FAT-PET for the estimates

	The FAT-PET MRA		
Mean	**PET meta-average**	**FAT asymmetry test**	**R^2**
0.066 (14.6)	0.029 (3.5)	0.403 (5.2)	0.015

The parentheses contain t-ratios.

TABLE 2 The 22 controls used in the AEL, aid effectiveness literature

	Control variable	No. incl.		Control variable	No. incl.
1	Aid *x* institutions	70	12	Size of government	250
2	Aid *x* policy	411	13	Regional dummies	789
3	Aid squared	333	14	Ethno-linguistic index	605
4	Aid lagged	463	15	Financial development	731
5	Capital controls	483	16	Trade openness	740
6	Human capital	238	17	Population size	292
7	Foreign direct investment	224	18	Income	1,274
8	Policies	530	19	2 eq. growth savings	44
9	Aid instability	815	20	2 eq. growth aid	58
10	Inflation	644	21	OLS	1,000
11	Fiscal stance	409	22	Africa	1,535

Some of the controls need an explanation: (1) Aid interacted with an institutional variable. (2) Aid interacted with a measure for good policies. (8) Some term for quality of policy. (14) Index for ethno-linguistic diversity of population. (15) Measure for financial deepening, such as bank balances over GDP. (18) GDP per capita is often in logs. (19) and (20) are two-equation models, with either a growth and a saving equation, or a growth and an aid equation. (21) Most non-OLS regressions try to account for simultaneity. (22) Some estimates are from Africa only. Here *Africa* is coded blank. See Doucouliagos and Paldam (Doucouliagos & Paldam, 2008, 2011) for more details on the coding.

4.2 | From meta-analysis to meta-mining: Augmenting the FAT-PET

The case study looks at 22 control variables that are all assumed to be ad hoc controls. These controls are listed in Table 2, which also shows how often they are included in the estimates. Two of these variables are conceptually different from the rest. It is the Africa-dummy, which is a *cp* control, and the OLS dummy, which gives a basic classification of the estimator.

Many researchers have taken the relation to be simultaneous, so it should be estimated by regression adjusting for that.[10] The first half of the studies contains few such regressions, but later studies have many. However, the results reached with OLS and more advanced methods do not differ (Net Appendix).

Table 3 shows what happens when the FAT-PET is augmented with 1 to 5 controls using the inclusion variable z for the variables of Table 2. The PET increases, and the FAT decreases. The augmentations produce a range of results reported in rows (3) to (7) of the table. Row (1) covers the 1,779 primary estimates displayed as the funnel of Figure 5. The meta-results from row (2) onwards differ much less than the primary estimates. If we compare rows (1) and (5), N is in the same order, but the std. is 7 times smaller in (5) than in (1). While the range of primary estimates in row (1) goes from −0.948 to 0.908, it only goes from −0.007 to 0.156 in row (7). Thus, all meta-results substantially reduce the range of primary results, but the augmentations still leave a wide range of choices.

I think that the profession believes that the partial correlation between aid and growth must be in the *reasonable range* from −0.02 to 0.10. It is reached with three augmentations. If we accept augmentations even when the literature has a publication bias, it gives a range of choices that include the full range of reasonable choices. In the same way, augmentations decrease the FAT, as seen in Figure 6b. The FAT is still positive, but it becomes insignificant in 32% of the cases and only rarely exceeds the basic test value. Thus, augmentations blunt the tool showing publication bias as predicted.

4.3 | Meta-mining through augmentations for 1 to 5 controls

Figure 7 shows the (FAT, PET)-scatter with all 26,334 combinations of five controls. A researcher who disregards the basic result and augments will obtain this selection of choices where preferences and interests can come into play.

The extreme ends of the scatter are highlighted as the black NW and SE points, each having 263 (1%) observations. To show aid ineffectiveness, your choice is the NW triangles, where the PET is about 0.011 and FAT is 0.406.

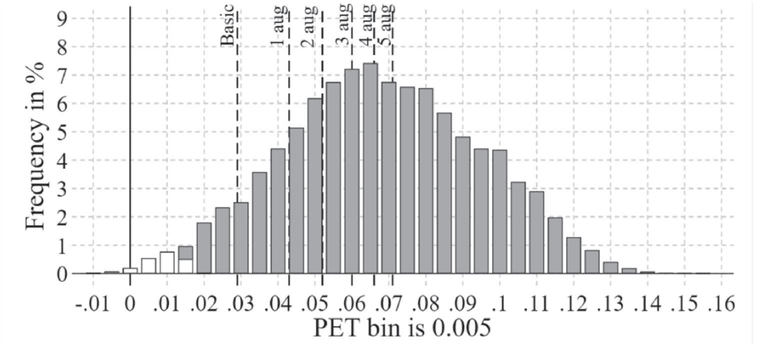
TABLE 3 Analyzing the 1,779 estimates of aid effectiveness

Estimates		N	Avr	Std	Min	Max	% ins.	Avr	Std.	Min	Max	% ins.
(1)	Primary	1,779	0.066	0.191	−0.948	0.908						
	Meta		PET meta-average					FAT asymmetry test				
(2)	Basic	1	**0.029**				0	**0.403**				0
(3)	1 aug	22	**0.043**	0.023	0.012	0.099	9.09	**0.341**	0.079	0.146	0.448	4.55
(4)	2 aug	231	**0.052**	0.026	0.002	0.117	7.79	**0.292**	0.094	0.036	0.481	10.39
(5)	3 aug	1,540	**0.060**	0.027	−0.005	0.130	5.00	**0.256**	0.096	−0.005	0.483	16.30
(6)	4 aug	7,315	**0.066**	0.027	−0.006	0.146	3.17	**0.227**	0.096	−0.040	0.483	24.90
(7)	5 aug	26,334	**0.071**	0.027	−0.007	0.156	2.11	**0.203**	0.094	−0.074	0.483	31.73

The table uses two abbreviations: *aug* means augmentations, and *ins* means insignificant. The Std is calculated across the N estimates. The number of possible choices is $N = \binom{22}{n}$ that rises as shown in the N-column.

[10]Doucouliagos and Paldam (2011) specifically analyze the effect of simultaneity controls, concluding they did not matter.

(a)



(b)

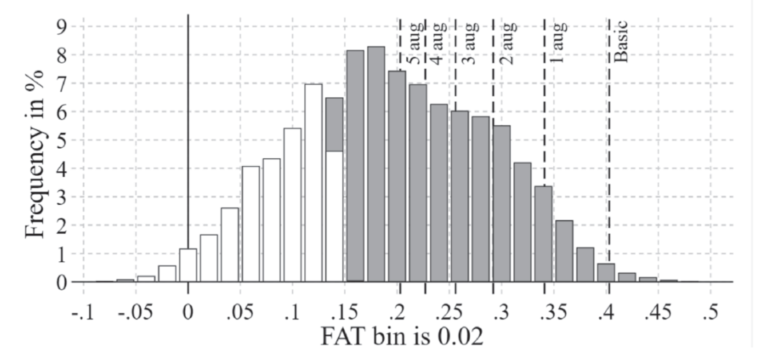


FIGURE 6 (a) The 26,334 PETs augmented with five controls (b) All 26,334 FATs augmented with five controls. Insignificant estimates are white, while significant ones are gray. Dashed vertical lines are average FATs from Table 3. The abbreviation *aug* means augmentation. See Paldam (2022b) on the distribution of the FATS and PETs.

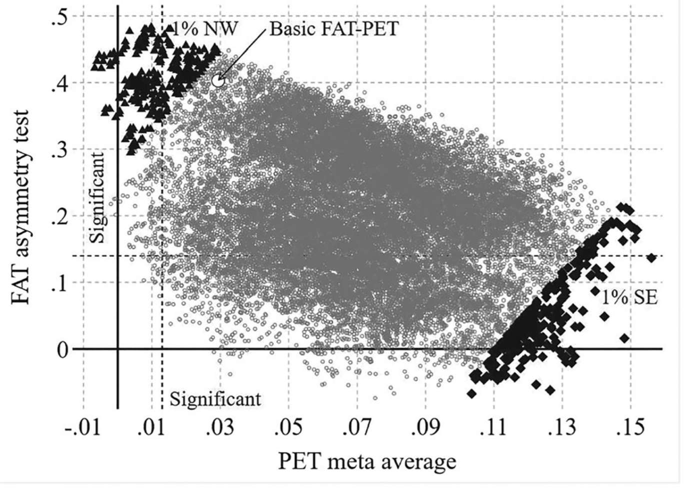


FIGURE 7 The (FAT, PET)-scatter of the 26,334 estimates with five augmentations

To show aid effectiveness, your choice is the SE diamonds, where the average PET is about 0.124 and the FAT is 0.048. This value of PET is almost twice the mean, and thus four times the basic meta-average, b_M.

The two dashed lines give the approximate division between significant and insignificant estimates according to Figure 6. The dashed vertical line for the PET and the dashed horizontal line for the FAT. The two sets of 263 larger black points are an assessment of the 1% extreme points to either side.

The average PET and FAT change along smooth curves, so they are easy to project; see the Net-Appendix. The projection says that for six augmentations, the basic FAT-PET is fully within the 1% NW set, and for nine augmentations, the average FAT is insignificant. Meta-mining has a wide scope. This has been used in practice:

The SE points include the choice of Mekasha and Tarp (2013).[11] They start by finding the same basic results as Doucouliagos and Paldam *op cit*. Then they augment and choose carefully, and conclude that aid is effective, and that the AEL is unbiased. Their paper does not suggest that they have selected an extreme end of a wide spectrum.

Thus, meta-analysis can be misused. If you are free to choose and you only show carefully selected augmented FAT-PETs, you can show what you want within a wide range. However, when the basic FAT-PET shows a publication bias, you are not free to choose. You should demonstrate that each control you use to augment is randomly included in the primary studies. If it is not, it should not be used for augmentation!

4.4 | The replication of the two most cited models

The data coded for a meta-study also allow the analyst to see if models with a new variable replicate. The most cited studies in the AEL in this century are *the good policy model* and the *aid squared model*.[12] Both papers introduced a new control variable that had a fine effect in the data chosen but proved hard to replicate outside that dataset.

The good policy model of Burnside and Dollar (2000) uses the interacted *Aid x policy* variable as the new control. It is variable 2 in Table 2. The variable has been included in 411 estimates. Tables A2 and A3 (in the Net Appendix) show that the variable does not work. On average, the estimates including the variable give lower aid effectiveness estimates than the ones without this variable. The two augmented PETs both suggest that the variable works (a little), but when the estimates of the authors are excluded, the replication fails.

The model of Hansen and Tarp (2000) has *aid squared* as the new control. It is variable 3 in Table 2. The variable has been used in 333 estimates in the literature. Here the estimates with and without the variable differ, but the augmented PETs show a negative effect of aid squared on aid effectiveness, see Tables A2 and A3 (in the Net Appendix).[13]

Thus, both models have failed at replication, and they have disappeared from the literature. This illustrates why results should be repeatedly replicated before they are trusted, and how meta-analysis allows replication of both the central model and specific model variants.

5 | CONCLUSION

Meta-analysis in economics is made to summarize the set of papers presenting regression results that claim to estimate the same parameter. It analyzes papers using the classic research strategy – or at least papers that are presented as if it was followed. Such papers start with a small literature survey showing why the paper presents something new, then follows a theory, leading to a model, which is operationalized as an estimation model. After a

[11] The skew reporting is likely to be due to strong preferences and interests of the authors, who are members of a research group at Copenhagen University, the DERG that is largely financed by aid – mainly from the Danish aid agency Danida. Mekasha and Tarp (2019) is an update without augmenting. It neatly replicates Doucouliagos and Paldam (2015), with only a few polemic remarks.

[12] The papers have 6,400 and about 2,800 citations in Google Scholar (Dec. 1st, 2021). The good policy researchers are from the World Bank, while the aid squared researchers are members of DERG (see the previous note).

[13] Once again, the results are even worse when the results of the authors are excluded. When the model failed, the authors made another model (see Dalgaard et al., 2004). The new model is different, but it has the same policy implications, as predicted by the discussion of Mekasha and Tarp *op cit*.

brief presentation of the data, the model is estimated, and in the great majority of cases the paper concludes that the theory is confirmed, i.e., it is not rejected.

It has often been shown that this strategy contains a great deal of make-believe, as it allows priors to play a large role. Researchers can – and often do – make many regressions and choose the best. Thus, preferences and interests have a wide scope. Figure 7 shows the amazing range of results reached by the literature on one well-defined parameter. Replication studies are surely important. Here I wish to add that the data used for the present paper are available from the author upon reasonable request.

Meta-analysis is a method to replace strict replication and show what the literature has found. It is important that the result of the meta-study is manipulation robust. The basic FAT-PET MRA is robust. The full literature should be collected, and a list published, so that any reader can check that it is an unbiased list, and the literature should be coded. It is some effort to find the literature and a major effort to do the coding. A few random coding errors are inevitable, but they matter little for the result. Once all of this is done, the basic result follows.

If the basic result shows a publication bias, one should only use augmented FAT-PETS for controls that are randomly included, as regards their effect on β. It is quite difficult to show that this is the case. If the FAT-PET is augmented with the variables generating the publication bias, it brings back the bias. Hence, it defeats the very purpose of meta-analysis.

ACKNOWLEDGMENTS

The paper was presented at the MAER-net colloquium in Piraeus (Greece) 2021, the Danish Public Choice Workshop in Aarhus 2022, and the European Public Choice 2022 Meeting in Braga (Portugal). I am grateful to the discussants, notably Bruno Frey. In addition, the paper owes much to discussions with Chris Doucouliagos, and comments from Erich Gundlach. I also wish to thank the referees for constructive advice. Mathias Barding has been an excellent research assistant.

DATA AVAILABILITY STATEMENT

The data used for the present paper are available from the author upon reasonable request.

REFERENCES

Burnside, C., & Dollar, D. (2000). Aid, policies and growth. *American Economic Review*, *90*, 847–868.

Dalgaard, C.-J., Hansen, H., & Tarp, F. (2004). On the empirics of foreign aid and growth. *The Economic Journal*, *114*(496), F191–F216.

Doucouliagos, H., & Paldam, M. (2006). Aid effectiveness on accumulation. A meta study. *Kyklos*, *59*, 227–254.

Doucouliagos, H., & Paldam, M. (2008). Aid effectiveness on growth. A meta study. *European Journal of Political Economy*, *24*, 1–24.

Doucouliagos, H., & Paldam, M. (2011). The ineffectiveness of development aid on growth: An update. *European Journal of Political Economy*, *27*, 399–404.

Doucouliagos, H., & Paldam, M. (2013). The robust result in meta-analysis of aid effectiveness: A response to Mekasha and Tarp. *Journal of Development Studies*, *49*, 584–587.

Doucouliagos, H., & Paldam, M. (2015). Finally a breakthrough? The recent rise in the size of the estimates of aid effectiveness, Cpt. 20. In M. Arvin & B. Lew (Eds.), *Handbook on the Economics of Foreign Aid* (pp. 325–349). Edward Elgar.

Doucouliagos, H., Paldam, M., & Stanley, T. D. (2018). Skating on thin evidence: Implication for public policy. *European Journal of Political Economy*, *54*, 16–25.

Hansen, H., & Tarp, F. (2000). Aid effectiveness disputed. *Journal of International Development*, *12*, 375–339.

Ioannidis, J. P. A., Stanley, T. D., & Doucouliagos, H. (2017). The power of bias in economics research. *Economic Journal*, *127*, 236–265.

Mekasha, T. J., & Tarp, F. (2013). Aid and growth. What meta-analysis reveals. *Journal of Development Studies*, *49*, 464–483.

Mekasha, T. J., & Tarp, F. (2019). A meta-analysis of aid effectiveness. Revisiting the evidence. *Politics and Governance*, *7*(2), 5–28.

Mueller-Langer, F., Fecher, B., Harhoff, D., & Wagner, G. G. (2019). Replication studies in economics - How many and which papers are chosen for replication, and why? *Research Policy*, *48*(1), 62–83.

Niskanen, W. A. Jr. (1994). *Bureaucracy and Public Economics (6 older essays)*. Edward Elgar.

Paldam, M. (2018). A model of the representative economist, as researcher and policy advisor. *European Journal of Political Economy*, *54*, 1–15.
Paldam, M. (2021). *The Grand Pattern of Development and the Transition of Institutions*. Cambridge UP.
Paldam, M. (2022a). Net-Appendix to section 4, available from http://martin.paldam.dk/Aid.php
Paldam, M. (2022b). Three sets of evidence about aid effectiveness. A survey. P.t., working paper available from http://martin.paldam.dk/Aid.php
Stanley, T. D., & Doucouliagos, H. (2012). *Meta-regression Analysis in Economics and Business*. Routledge.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Paldam, M. (2022). Meta-mining: The political economy of meta-analysis. *Kyklos*, 1–16. https://doi.org/10.1111/kykl.12321

APPENDIX: The logic of the FAT-PET. The variables and assumptions

The β-literature gives N observations of $(b, s) = (b, 1/p)$, where p is the precision. The *funnel* is the scatter diagram (b, p). It is broad at the base for low precision and narrows for high precision. Figure 5 shows a typical funnel. Figures 2, 3 and 4 are stylized funnels. Statistical theory and many simulations show that model uncertainty should generate symmetric funnels, where the axis of symmetry is the mean. However, most funnels have asymmetries that can be explained as censoring of 'bad' results by the priors of the researcher as explained in the main text. Obviously, there are more estimates to censor at low precision.

The FAT-PET MRA is the regression (1a) $b = b_M + b_F s$ or (1b) $b = b_M + b_F/p$

The b_M is the PET, for Precision (weighted) Estimate Test.[14] The b_F is the FAT, for Funnel Asymmetry Test. It is an MRA, Meta Regression Analysis, as it is a regression on regression coefficients (the bs); see Stanley and Doucouliagos (2012).

The intuition behind the FAT-PET is easy to grasp from version (1b). It shows that the MRA converges to b_M for p going to infinity. Thus, b_M is an estimate of the meta-average. If the funnel is symmetric $b_F \approx 0$.

Table. Variables, terms, assumptions, and the meta-tools used.

Variables			
(1)	β	Parameter of interest	Estimated in the β-literature
(2)	F	Theoretical model	$g = F(h, \ldots)$, where $\beta = \partial g/\partial h$
(3)		$g = \alpha + \beta h + \gamma\zeta + []$	Estimating equation, with controls ζ ($=\zeta_1$) and $[] = \gamma_2\zeta_2 + \ldots + \gamma_n\zeta_k$
(4)	ζ	Control discussed	ζ is not in the theory. It is included in some but not all estimates
(5)	z	Inclusion variable	Binary (0,1) variable. It is 1 if ζ is included in the estimate, else 0
(6)	b_z	The effect of ζ on b	The difference of b if the estimates is run with and without ζ
The β-literature consists of N primary papers			
(7)	b_i	Estimate $b_i \approx \beta$	***Fact:*** most insignificant or negative estimates remain unpublished
(8)	p_i	Precision of b_i	$p = 1/s = t/b$, where s is standard error and t is t-ratio
(9)	g_i	Estimate $g_i \approx \gamma$	***Fact:*** if g_i is insignificant, control ζ is normally omitted
Statistics for the whole of the β-literature of $i = 1, \ldots, N$ papers			
(10)	$\bar{b}$	Mean of estimates	$\bar{b} > 0$, in accordance with theory
(11)	b_M	Meta-average (PET)	Estimate of β adjusted for bias to be closer to β than is $\bar{b}$
Matrices with one row for each paper B, Z and A			
(12)	B	$(N \times 2)$-matrix	Row i is (b_i, p_i)
(13)	Z	(N x K)-matrix	Element z_{ij} in Z is one if control ζ_j is included in study i, else zero
(14)	A	(N x L)-matrix	Other characteristics of paper, such as author, journal, etc., not discussed

The rows in Z indicate how each estimate is reached.

[14]The meta-average may also be estimated by WLS (Weighted Least Squares), where the precisions are used as weights. The two estimates are typically rather close. The present only uses the PET.