Aarhus 7/8-2013

Net Appendix to:

How do partly omitted control variables influence the averages used in meta-analysis in economics?

Martin Paldam,[1] Department of Economics and Business, Aarhus University, Denmark

Content:

Section		Page
1	Introduction Table 1	2
2	The ideal funnel and censoring Table 2 and 7 Figures 2a to 2g	4
3	Experiments with the correlation ρ and the effect of POC1 Table 3 and 5 Figures 3a to 3e	5
4	Experiments with the inclusion frequency q and the effect of POC2 Table 4 and 5 Figures 4a to 42	12
5	Combinations of DGPs and EMs: First set of experiments Table 5 and 9 Figures 5a and 5i	15
6	Censoring the cases with POCs Table 6 and 9 Figures 6a to 6i	16
7	Combinations of DGPs and EMs: Second set of experiments Table 7 and 9 Figures 7a and 7i	21
8	The simulation program Tables 8a and 8b	24
9	Conclusions	26

Acknowledgment and notes: The simulation program used is written by Jan Ditzen. This paper contains no references – they are given in the main paper. For easy cross reference the table (tables) and the figures in each section has the same number as the section.

[1]. Street address: Fuglesangs Allé 4, Building 2632 (L), DK-8210 Aarhus V. E-mail: mpaldam@econ.au.dk.
URL: http://www.martin.paldam.dk, for project add /Mpapers-method.php.

1. Introduction

The main paper studies some problems by simulation experiments. However, many more simulations have been needed to study these problems. This appendix documents all 182 simulation experiments made. Each experiment needs *N x R* regressions, which is *N* "primary" regressions for one funnel, and *R* funnels reach a precise average. This is done with a set of programs outlined in section 8 below. They have three loops within each other: An inner regression loop, a middle funnel loop and an outer experiment loop.

One funnel is a representation of the distribution of *N* simulated average estimates of the parameter $\beta = 0.25$. The averages over one funnel are $\underline{b}$, β_M, β_{AR} and β_{AW}. They are defined in the paper and in Table 1. Over a full experiment the averages are $\underline{b}_a$, $\underline{\beta_M}$, $\underline{\beta_{AR}}$ and $\underline{\beta_{AW}}$. Since we know the true value, we can study the fit of the 4 average estimates.

Table 1. The DGP and EM, and the main parameters

The (DGP, EM) pair, x is variable of interest z_1 and z_2 are the POCs and ε is noise				
	DGP, data generating process	EM, estimating model	Variables	
(1)	$y_i = \beta x_i + \varepsilon_i$	$y_i = b_i x_i + u_i$	y_i	Dependent variable
(2a)	$y_i = \beta x_i + \gamma_1 z_{1i} + \varepsilon_i$	$y_i = b_i x_i + g_1 \omega_{1i} z_{1i} + u_i$	x_i	Variable of interest
(2b)	$y_i = \beta x_i + \gamma_2 z_{2i} + \varepsilon_i$	$y_i = b_i x_i + g_2 \omega_{2i} z_{2i} + u_i$	z_{1i}	POC1, control variable
(3)	$y_i = \beta x_i + \gamma_1 z_{1i} + \gamma_2 z_{2i} + \varepsilon_i$	$y_i = b_i x_i + g_1 \omega_{1i} z_{1i} + g_2 \omega_{2i} z_{2i} + u_i$	z_{2i}	POC2, control variable

The parameters: β is fixed at 0.25, while γ_1, γ_2, ρ and q are varied			
β	Parameter of interest $\beta = 0.25$	b_i	Estimate of β
ε	Generated noise term	u_i	Estimated residual
γ_1	Effect of POC1, z_1, in DGP	g_{i1}	Estimate of γ_1
γ_2	Effect of POC2, z_2, in DGP	g_{i2}	Estimate of γ_2
q	Inclusion probability for the POCs	ω, φ	Binary inclusion/exclusion vector $\omega_i + \varphi_i = 1$
ρ	Correlation of x and the POCs. When $\rho \neq 0$ POC biases occurs		

Averages from			Conventions used in tables	
Funnel	Exprm.		R	Number of funnels in each experiment
$\underline{b}$	$\underline{b}_a$	Arithmetic mean	N	Number of regressions run for the funnel
β_M	$\underline{\beta_M}$	PET, meta average	M	Sample size for estimate, $M = 20, 21, \ldots, 19 + N$
β_{AR}	$\underline{\beta_{AR}}$	Rightly augmented m.a.	*Avs* [a)]	Count of not rejected estimates of $\beta = 0.25$ at 5 %
β_{AW}	$\underline{\beta_{AW}}$	Wrongly augmented m.a.	*Fs*	Count of FATs showing no asymmetry at 5 % level
Averages within 5% of the true β are bolded. This is estimates between and 0.237 and 0.263				

Note: The equations have no constant, so the exogenous variables, x, z_1, z_2 and the noise term ε have zero means.
a. The same terminology applies to β_{AE} % and β_{AW} %.

The simulations of each primary regression use a (DGP, EM)-pairs, where the DGP is the *data generating process*, and EM is the *estimating model.* Table 1 gives the four DGPs and EMs used. Equation (1) is used in section 2 only. Sections 3 and 4 use the same (DGE, EM)-pair and vary the four variable parameters: γ_1, γ_2, ρ and q. Sections 5 to 7 looks at the realistic situation where the researcher does not know the DGP, so that the (DGE, EM)-pair often differ. Here only the possibilities (2a), (2b) and (3) are permitted. This gives 3 x 3 = 9 combinations, where only 3 have the same DGP and EM.

All results are from simulated funnels with N = 500 points, where each is estimated by a regression on simulated data. The main paper reports average results from R = 1,000 simulated funnels per line in the tables. Although high Rs gives a high precision, it is not necessary to see the pattern in the results, when the parameters are varied, so all lines in the tables below are average results from only R = 100 simulated funnels. This means that each line is estimated from 100 x 500 = 50'000 regressions. The average estimate is bolded if it is within 5% of the true value β = 0.250, so bolded averages are between 0.237 and 0.263.

The three averages β_M, β_{AR} and β_{AW} are estimated, so they come with a t-ratio and with a funnel asymmetry test FAT, β_F, β_{FR} and β_{FW} respectively. For each pair (β_x, β_{Fx}) it is tested if $\beta_x \approx \beta$ and $\beta_{Fx} \approx 0$ at the 5 % level of significance, by the two counts Avs_x and Fs_x, which are both a number between zero and R. The results for each experiment are 10 statistics: ($\underline{b_a}$, $\underline{\beta_M}$, Avs_1, Fs_1, $\underline{\beta_{AR}}$, Avs_2, Fs_2, $\underline{\beta_{AW}}$, Avs_3, Fs_3). They are presented as a row in a table. Some of these rows are illustrated by a graph typically covering one funnel allowing the readers (and the author) to see what is going on. By using R = 100 the two counts become becomes integers in %. As we are using the 5 % level of significance throughout, a result is fine if $Avs \approx Fs \approx 95$. For most bolded averages Avs and Fs are about 95. They are often a bit lower, pointing to the power of the tests.

The N regressions for each funnel are made on simulated data with sample size M = 20, 21, 22, …, N + 20. The changing sample size gives variation in the precision of the estimates, as needed for realistic funnels. When some estimates are censored and N is increased, the last regressions have a higher sample size. Thus, the top of the censored funnels have higher precisions the larger the censoring is. This is likely to be realistic.

Most funnels in models with one or two POCs have two tops. The augmentations make one or the other top go away. Since both POCs are generated by the same distribution, they both give the same FAT, when the other is deleted. This is confirmed by all tables below. Thus, the fact that the FAT is fine only means that a nice symmetrical funnel has been generated. It might be the right or the wrong one.

2. The ideal funnel and censoring

The first simulations looks at the case where DGP = EM = 1, so formally the model contains no POCs. This is taken to cover a situation with many randomly included POCs. The variation in the estimates is generated by ε, and the different sample sizes. The results are reported in Table 2. Since there is no (formal) POCs there is only one possible meta average. It has two panels:

Table 2. The ideal funnel in row (1) and 11 cases of censoring

Panel A: Cases of censoring with ε = 2: Censoring is: none, -0.2, -0.15, … , 0.40															
Row	Censo-ring	N	Averages		Accept true		Row	Censo-Ring	N	Averages		Accept true	
			b_a	β_M	*Avs*	*Fs*				b_a	β_M	*Avs*	*Fs*
(1)	None	500	**0.250**	**0.248**	92	89	(8)	0.10	850	0.275	0.212	2	0
(2)	-0.20	550	**0.253**	**0.246**	93	89	(9)	0.15	900	0.287	0.212	0	0
(3)	-0.15	600	**0.258**	0.235	81	78	(10)	0.20	950	0.308	0.225	12	0
(4)	-0.10	650	**0.258**	0.232	75	72	(11)	0.25	1000	0.338	**0.250**	94	0
(5)	-0.05	700	**0.261**	0.227	61	41	(12)	0.30	1200	0.379	0.287	0	0
(6)	0	750	**0.263**	0.223	36	19	(13)	0.35	1700	0.427	0.331	0	0
(7)	0.05	800	0.268	0.217	12	02	(14)	0.40	2500	0.497	0.374	0	0

Panel B: Cases of censoring at 0.25, for N = 1000, with ε = 0.5, 1, 2, 3, 4, 5, 6 and 7													
Row	Noise	Averages			Accept true		Row	Noise	Averages			Accept true	
	ε	b_a	β_M	b_a/β_M	*Avs*	Fs		ε	b_a	β_M	b_a/β_M	*Avs*	*Fs*
(1)	0.5	0.272	**0.250**	1.088	94	0	(5)	4	0.427	**0.251**	1.703	94	0
(2)	1	0.294	**0.250**	1.176	94	0	(6)	5	0.471	**0.251**	1.878	94	0
(3)	2	0.338	**0.250**	1.352	94	0	(7)	6	0.515	**0.251**	2.053	94	0
(4)	3	0.383	**0.251**	1.528	94	0	(8)	7	0.560	**0.251**	2.228	94	0

Note: N includes the censored, so N has to be higher than 500 to give 500 funnel points, in rows (2) to (14) in Panel A, and in all experiments of Panel B. Recall that averages within 5 % of the true value 0.250 are bolded.

Figure 2. Illustrations of the funnels in Panel A of Table 2

Figure 2a. Line (1): The ideal funnel   Figure 2b. Line (6): Censoring at zero

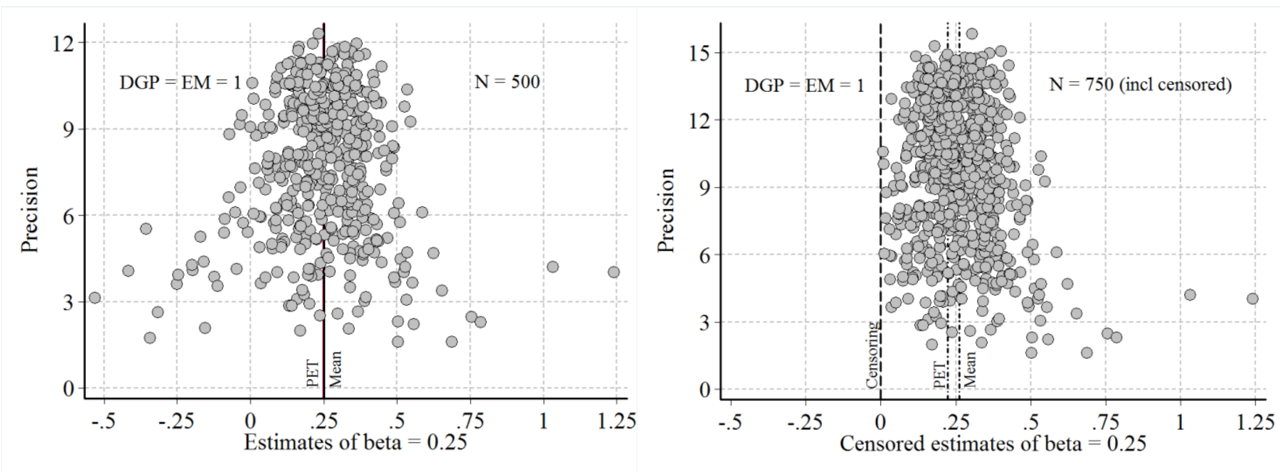


Figure 2c. Line (9): Censoring at 0.15

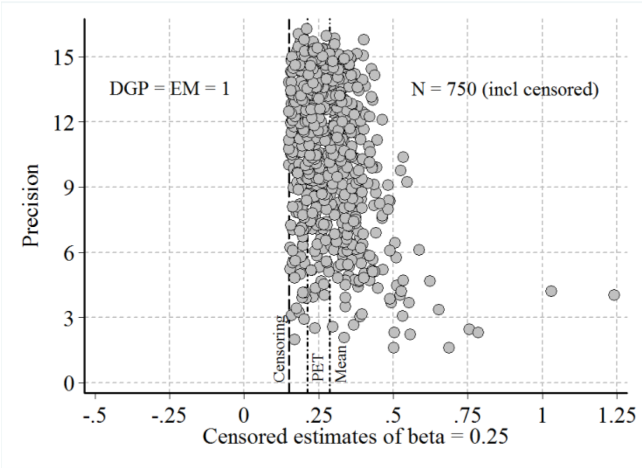


Figure 2d: Line (11): Censoring at 0.25

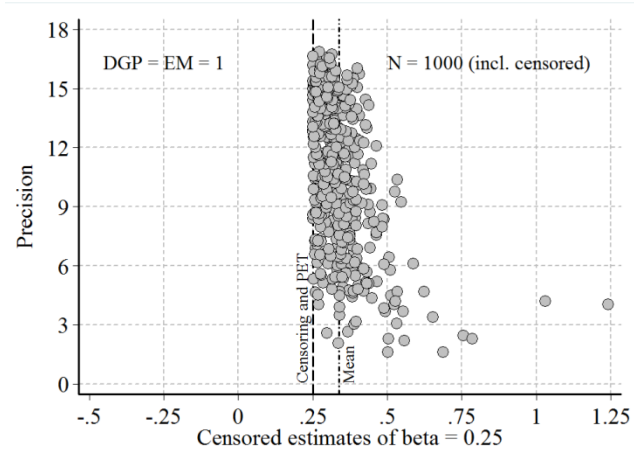


Figure 2e. Line (13): Censoring at 0.35

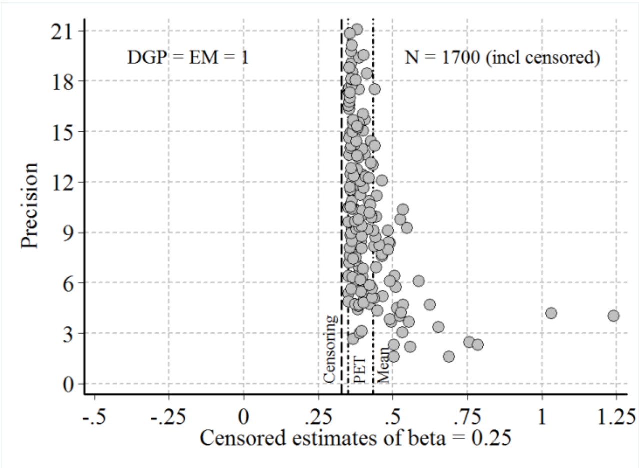
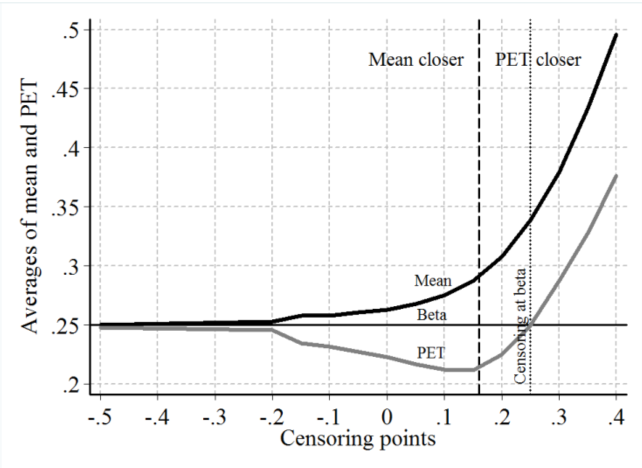


Figure 2f. The bias of the mean and the PET



Panel A: Row (1) is the ideal case, with no censoring. It has virtually the same mean $\underline{b}$ = 0.250 and PET β_M = 0.248. Figure 2a shows a specimen of an ideal funnel with these parameters. It is symmetrical and has one peak.

The mean and the PET-average, β_M, included on the figure are from Table 2. Thus, it is more precise than the estimates given on the funnel. This practice is used in all funnels shown.

Rows (2) to (14) report gradually more censored versions of the funnel: This is taken to cover the case where the random inclusion of the POCs becomes increasingly systematic. To get approximately 500 points in the funnel, when some points are censored, N is increased – this causes the number of observations in the regressions to increase so precision increases.

As long as the censoring is moderate the mean may still be closer to β = 0.25. But as the censoring increases the mean moves away from β, while the PET meta-average becomes better. The point where the PET becomes the best average is about halfway between zero and β (= 0.25). The

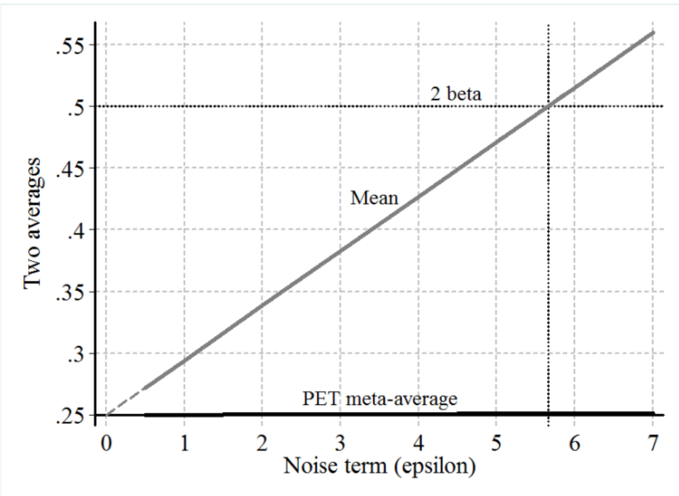
PET is perfect, when the censoring is in the middle of the distribution, i.e., at β. For censoring above β both averages quickly turns bad, but the PET is still the best. It is no wonder that it is difficult to catch the true average in such cases of extreme censoring, see Figure 2e, where there is nothing left of the peak for the PET to converge too.

Figure 2f shows the two averages as functions of the point of censoring. It only uses the averages reported in Table 2, so the curves are not perfectly smooth. From the small "kinks" on the curve we conclude that the estimates based on 100 funnels have an uncertainty of $\pm$ 0.01, which is fine for the purpose of this appendix.

Panel b: The reader will note that when the censoring at exactly β = 0.25, gives a PET meta-average of 0.250 that is precisely right. The panel show that this generalizes for all values of ε.

If the economic theory and beliefs are right then surely the researcher wants to find an estimate that is at β or above. Thus, it is likely that the point of censoring is close to the true value for β, which is 0.25. Thus, the PET meta-average might be the perfect average.

Figure 2g. Illustration of the funnels in Panel B of Table 2



The results also show that the publication bias (defined as $\beta_M - \underline{b}$) is proportional to the width of the funnel as measured by ε. This is illustrated on figure 2g. The PET-line is the same as the 0.25 grid-line and the mean is a straight line that starts in 0.250 for ε = 0.

The folk theorem that the expected relative publication bias found in a meta-study is 2, so that the ratio $\underline{b}_a/\underline{\beta}_M = \underline{b}_a/\underline{\beta}_M \, \beta = 2$, so that $\underline{b}_a = 0.5$, occurs at approximately ε = 5.67, which is a rather wide funnel.

3. Experimenting with the correlation ρ and the effect of POC1

The present and the next section show a set of experiments with four variable parameters: γ_1, γ_2, ρ and q. To get as many possibilities both sections uses the (DGP, EM)-pair of (3, 3) where all four parameters are at play. As the two POCs have error terms of 1, the noise term $\varepsilon = 1$. This makes the funnels about as wide as the ones in section 2, but now most are two-topped due to the POC biases.

The present section fixes q and γ_2 at 0.5 and -0.5 respectably. Table 3a shows results, where ρ and γ_1 are varied. The correlation ρ between x and z_1 takes 6 values: ρ = 0.75, 0.5, 0.25, 0, -0.25 and -0.5. The effect, γ_1, of z_1 in the DGP takes 7 values: γ_1 = -0.75, -0.50, -0.25, 0, 0.25, 0.5 and 0.75. This gives the 6 *x* 7 lines in the table. Due to the results for the PET meta-average the range for γ_1 is extended to 2.5 in Table 3b.

The calculations have an *imposed symmetry*: q and ρ are the same. Hence, the two POCs cancel out, when $\gamma_1 = -\gamma_2$. This imposed symmetry is used to check the calculations. In the case of the present section γ_2 is fixed at -0.5, so it happens when γ_2 is -0.5 as well. Note that in the cases where this condition holds all averages are at the true value $\beta = 0.25$. This gives a line with bolded averages for each value of ρ. The table has five sections as indicated by the shading. Section 1 gives the values of the two parameters. The remaining sections report an average:

The first average in (3) is the (arithmetic) mean. Then follow 2 augmented meta averages: The rightly augmented in (4) and the wrongly augmented in (7). As $q = 0.5$ we know from the main paper that *when there is no censoring* the mean is the average of the rightly and the wrongly augmented averages. This holds rather well. Finally, (10) is the PET meta-average. It is typically the worst of the four averages, as discussed in the paper. Also, note that when $\rho = 0$, the POCs are irrelevant for the term of interest in the DGP. It only matters what x is. Hence, all estimates of β when $\rho = 0$, are true, and hence bolded.

The EM is $y_i = b_i x_i + g_1 \omega_{1i} z_{1i} + g_2 \varphi_{2i} z_{2i} + \varepsilon_i$. We want the meta-average to include the POCs in the cases where they are excluded. Hence, when I use to use the exclusion vectors φ_1 and ω_2 in the right augmentation and the reverse augmentation is the wrong one.

The main pattern in the numbers calculated is shown on Figures 3a to 3d. They show the bias, which is the average minus 0.250, for the 6 values of ρ. Due to the symmetries imposed all six ρ-lines on the 4 graphs intersect at zero, and the line for $\rho = 0$ is horizontal.

Figure 3c shows that the right augmentation works rather well.

Table 3a. Experiments with ρ and γ_1 – the main range

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
For	POC1	Mean	Right augmentation			Wrong augmentation			PET meta average		
ρ	γ_1	$\underline{b}_a$	$\underline{\beta}_{AR}$	*Avs*	*Fs*	$\underline{\beta}_{AW}$	*Avs*	*Fs*	$\underline{\beta}_M$	*Avs*	*Fs*
0.75	-0.75	0.156	**0.250**	94	96	0.062	0	96	-0.009	0	0
	-0.5	**0.250**	**0.250**	93	96	**0.250**	97	96	**0.249**	98	99
	-0.25	0.344	**0.250**	93	95	0.437	0	95	0.511	0	0
	0	0.438	**0.250**	94	95	0.625	0	95	0.774	0	0
	0.25	0.533	**0.251**	95	95	0.812	0	95	1.036	0	0
	0.5	0.627	**0.251**	95	95	1.000	0	95	1.294	0	0
	0.75	0.721	**0.251**	95	95	1.188	0	95	1.539	0	0
0.50	-0.75	0.187	**0.249**	95	95	0.124	0	95	0.179	0	97
	-0.50	**0.250**	**0.249**	94	95	**0.249**	97	95	**0.249**	96	96
	-0.25	0.313	**0.250**	94	95	0.374	0	95	0.365	0	4
	0	0.375	**0.250**	94	94	0.499	0	94	0.492	0	0
	0.25	0.438	**0.250**	96	95	0.624	0	95	0.597	0	0
	0.50	0.501	**0.250**	96	95	0.749	0	95	0.643	0	6
	0.75	0.563	**0.250**	95	94	0.874	0	94	0.601	0	97
0.25	-0.75	**0.219**	**0.249**	94	93	0.187	0	93	**0.253**	96	21
	-0.50	**0.250**	**0.250**	92	92	**0.250**	94	92	**0.250**	93	89
	-0.25	**0.282**	**0.250**	92	91	0.313	0	91	0.270	50	85
	0	0.313	**0.250**	92	91	0.376	0	91	0.300	4	88
	0.25	0.345	**0.250**	92	91	0.438	0	91	0.311	5	61
	0.50	0.376	**0.250**	92	93	0.501	0	93	0.288	58	0
	0.75	0.408	**0.250**	94	95	0.564	0	95	0.231	87	0
0	-0.75	**0.250**	**0.250**	91	90	**0.249**	92	90	**0.250**	89	88
	-0.50	**0.250**	**0.250**	90	89	**0.250**	92	89	**0.250**	89	88
	-0.25	**0.250**	**0.250**	88	89	**0.250**	92	89	**0.250**	90	88
	0	**0.250**	**0.250**	88	89	**0.250**	91	89	**0.250**	89	88
	0.25	**0.250**	**0.250**	90	90	**0.250**	90	90	**0.250**	90	88
	0.50	**0.250**	**0.250**	90	91	**0.250**	91	91	**0.249**	91	91
	0.75	**0.250**	**0.250**	91	92	**0.250**	92	92	**0.249**	93	92
-0.25	-0.75	**0.281**	**0.250**	97	95	0.312	0	95	**0.246**	98	18
	-0.50	**0.250**	**0.250**	97	94	**0.250**	94	94	**0.250**	96	96
	-0.25	**0.219**	**0.250**	95	94	0.187	0	94	0.230	53	79
	0	0.187	**0.250**	96	94	0.125	0	94	0.202	7	82
	0.25	0.156	**0.250**	96	93	0.062	0	93	0.191	6	55
	0.50	0.125	**0.250**	95	93	0.000	0	93	0.215	61	1
	0.75	0.100	**0.250**	96	92	-0.050	0	92	**0.259**	99	0
-0.50	-0.75	0.314	**0.251**	96	94	0.376	0	94	0.322	0	96
	-0.50	**0.251**	**0.251**	94	94	**0.251**	94	94	**0.251**	96	95
	-0.25	0.188	**0.251**	94	95	0.126	0	95	0.135	0	7
	0	0.125	**0.251**	95	93	0.001	0	93	0.007	0	0
	0.25	0.062	**0.251**	96	94	-0.124	0	94	-0.098	0	0
	0.50	-0.002	**0.252**	96	93	-0.249	0	93	-0.144	0	1
	0.75	-0.052	**0.252**	96	93	-0.348	0	93	-0.119	0	80

Note: DGP = EM = 3, β = 0.25, N = 500, R = 100, γ_2 = 0.5 and q = 0.5.

Table 3b. Experiments with ρ and γ_1 – extension of the table

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
For	POC1	Mean	Right augmentation			Wrong augmentation			PET meta average		
ρ	γ_1	$\underline{b_a}$	$\underline{\beta_{AR}}$	*Avs*	*Fs*	$\underline{\beta_{AW}}$	*Avs*	*Fs*	$\underline{\beta_M}$	*Avs*	*Fs*
0.75	0.80	0.739	**0.250**	95	94	1.225	0	94	1.585	0	0
	0.85	0.758	**0.250**	95	94	1.262	0	94	1.630	0	0
	0.90	0.777	**0.250**	94	90	1.300	0	94	1.674	0	0
	1	0.814	**0.250**	95	94	1.375	0	94	1.756	0	0
	1.5	1.002	**0.250**	96	95	1.750	0	95	1.994	0	0
	2	1.190	**0.250**	96	94	2.125	0	94	1.790	0	0
	2.5	1.378	**0.250**	96	94	2.501	0	94	1.136	0	77
0.50	0.80	0.576	**0.250**	94	93	0.899	0	93	0.581	0	100
	0.85	0.588	**0.250**	94	93	0.924	0	93	0.558	0	97
	0.90	0.601	**0.250**	94	93	0.949	0	93	0.531	0	85
	1	0.626	**0.250**	94	92	0.999	0	92	0.468	0	26
	1.5	0.751	**0.250**	94	90	1.249	0	90	0.087	22	0
	2	0.876	**0.250**	94	89	1.499	0	89	-0.200	0	0
	2.5	1.001	**0.250**	93	89	1.750	0	89	-0.350	0	0
0.25	0.80	0.413	**0.251**	95	95	0.577	0	95	0.218	77	0
	0.85	0.419	**0.251**	95	96	0.589	0	96	0.204	49	0
	0.90	0.426	**0.251**	95	96	0.602	0	96	0.190	21	0
	1	0.438	**0.251**	95	95	0.627	0	95	0.163	2	0
	1.5	0.501	**0.251**	96	97	0.753	0	97	0.056	0	0
	2	0.564	**0.252**	96	96	0.879	0	96	0.007	0	0
	2.5	0.626	**0.252**	95	94	1.005	0	94	-0.012	0	0
0	0.80	**0.250**	**0.250**	93	93	**0.250**	92	93	**0.250**	93	95
	0.85	**0.250**	**0.250**	93	93	**0.250**	92	93	**0.250**	93	95
	0.90	**0.250**	**0.250**	93	94	**0.250**	92	94	**0.250**	94	96
	1	**0.250**	**0.250**	92	94	**0.251**	94	94	**0.249**	94	96
	1.5	**0.250**	**0.250**	93	94	**0.251**	94	94	**0.249**	94	93
	2	**0.251**	**0.250**	94	94	**0.251**	95	94	**0.249**	94	93
	2.5	**0.251**	**0.250**	94	94	**0.251**	95	94	**0.249**	95	95
-0.25	0.80	0.087	0.251	95	93	-0.075	0	93	0.286	65	0
	0.85	0.081	0.251	95	93	-0.087	0	93	0.200	40	0
	0.90	0.074	0.251	95	93	-0.100	0	93	0.314	14	0
	1	0.062	**0.251**	95	93	-0.125	0	93	0.341	0	0
	1.5	-0.001	**0.251**	95	96	-0.249	0	96	0.447	0	0
	2	-0.063	**0.251**	97	96	-0.374	0	96	0.495	0	0
	2.5	-0.126	**0.251**	97	96	-0.499	0	96	0.514	0	0
-0.50	0.80	-0.077	0.252	95	94	-0.398	0	94	-0.084	0	98
	0.85	-0.090	0.252	95	94	-0.423	0	94	-0.061	0	98
	0.90	-0.103	0.252	95	94	-0.448	0	94	-0.034	0	88
	1	-0.128	**0.252**	95	93	-0.498	0	93	0.029	1	23
	1.5	-0.254	**0.253**	95	91	-0.747	0	91	0.410	15	0
	2	-0.380	**0.253**	95	94	-0.996	0	94	0.700	0	0
	2.5	-0.506	**0.253**	95	95	-1.245	0	95	0.851	0	0

Note: DGP = EM = 3, β = 0.25, N = 500, R = 100, γ_2 = 0.5 and q = 0.5.

Figure 3. The pattern of biases in Table 3: The deviations of the four averages from $\beta = 0.25$

Figure 3a. Bias of the mean

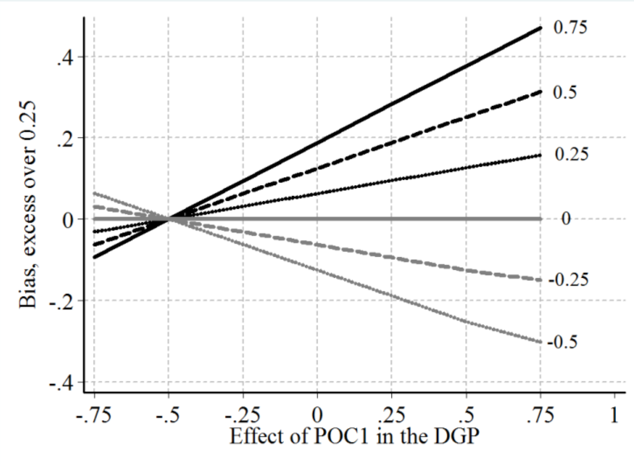


Figure 3b. Bias of β_M – see Figure 3e

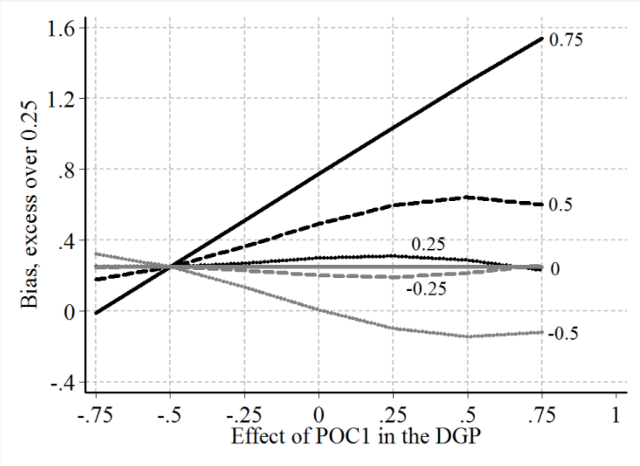


Figure 3c. Bias of β_{AR} (note scale)

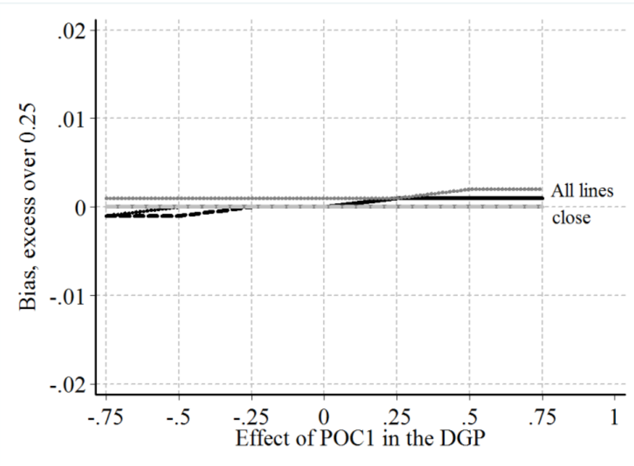


Figure 3d. Bias of β_{AW}

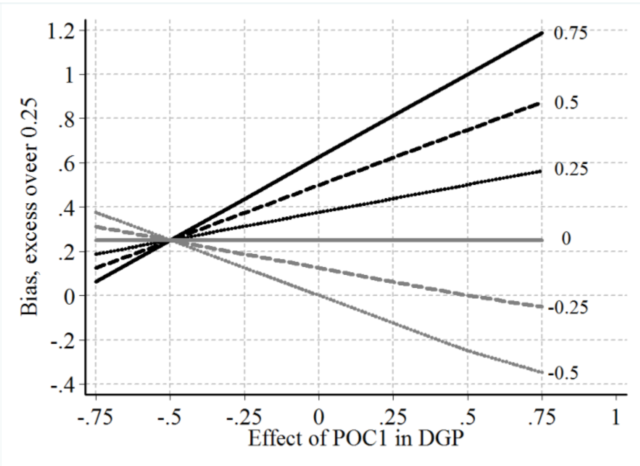
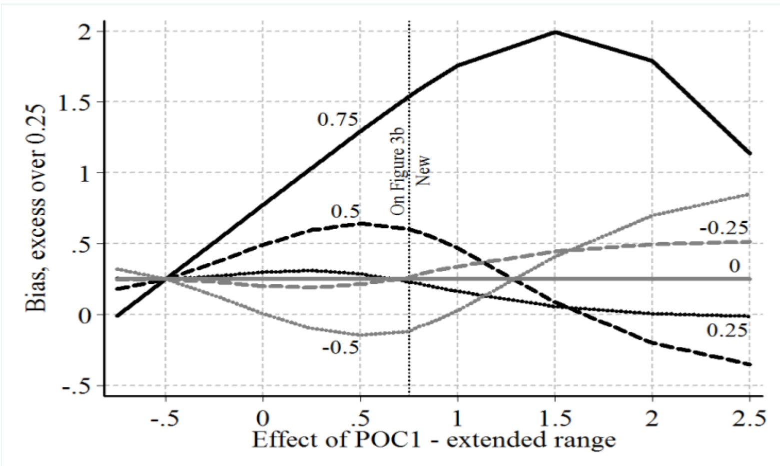


Figure 3e. Extended version of Figure 3b



Given the observations above, it is easy to interpret Figures 3d. Here the bias become worse the larger the deviation is from the imposed symmetry conditions, i.e., the more γ_1 deviate from -0.5 and the larger numerical value of ρ is. It is also easy to see that all points on Figure 3a are the average of the corresponding points on figures 3 c and 3 d.

The interesting curve to interpret is Figure 3b for the PET. Here the curves for ρ bend. This is why Table 3b was calculated covering the range where the bend occurs and the high end. The4 extended version of Figure 3b is shown as Figure 3e.

The Figure show that when the POC becomes really powerful in the DGP it comes to dominate in the relation.

4. Experiments with the inclusion frequency *q* and the effect of POC2

The present section fixes ρ and γ_1 at 0.5 and -0.5 respectably. Table 3 shows results, where q and γ_2 are varied. The inclusion probability for the POCs takes 6 values: q = 0.9, 0.8, 0.65, 0.35 and 0.2. The effect of z_2 in the DGP takes 7 values: γ_1 = -0.7, -0.5, -0.3, 0, 0.3, 0.5 and 0.7. This gives the 6 *x* 7 lines in the table.

The averages are shown on Figure 4 that is constructed as Figure 3. Once again the three of the graphs are easy to interpret; while Figure 4b shows that the PET meta-average bends, though in a more simple way than in Figure 3b. As q varies the weights of the two peaks in the mean varies, so even when the wrong augmentation in figure 4d has some lines that overlaps, it still give different lines in Figure 4a.

Figure 4. The pattern in Table 4: The deviations of the four averages from β = 0.25

Figure 4a. Bias of the mean

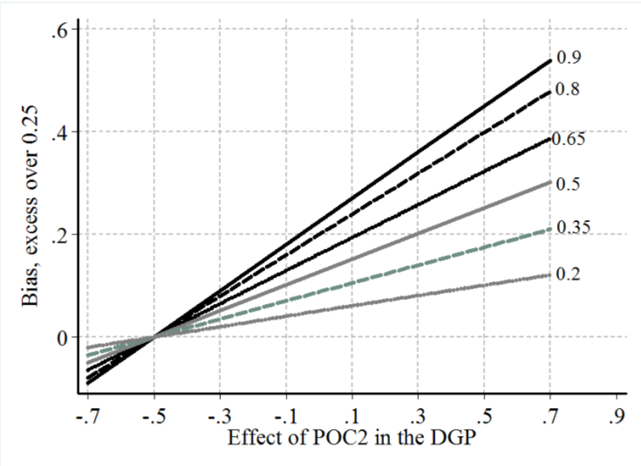


Figure 4b. Bias of β_M -- see Figure 4e

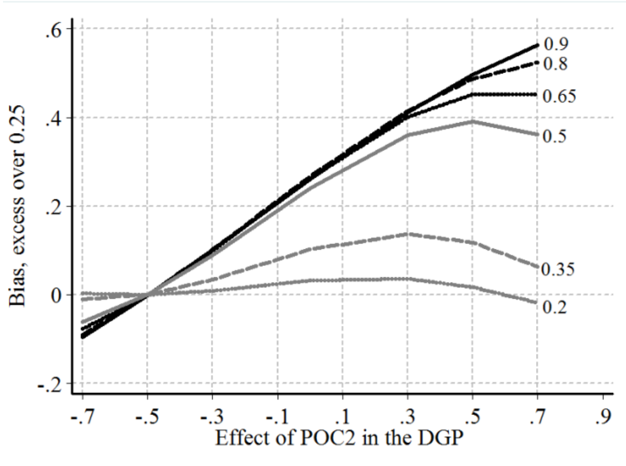


Figure 4c. Bias of β_{AR} (note scale)

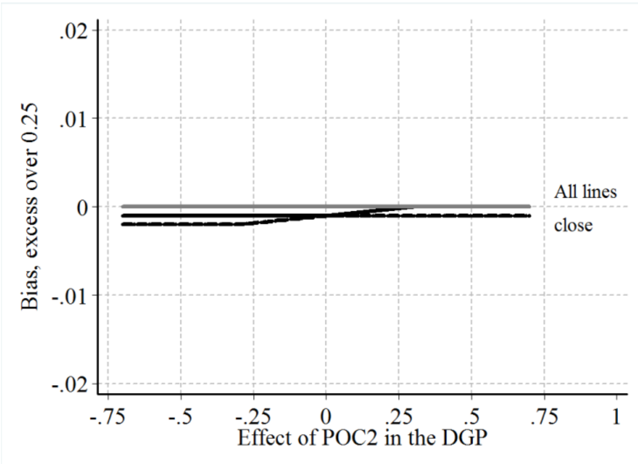


Figure 4d. Bias of β_{AW}

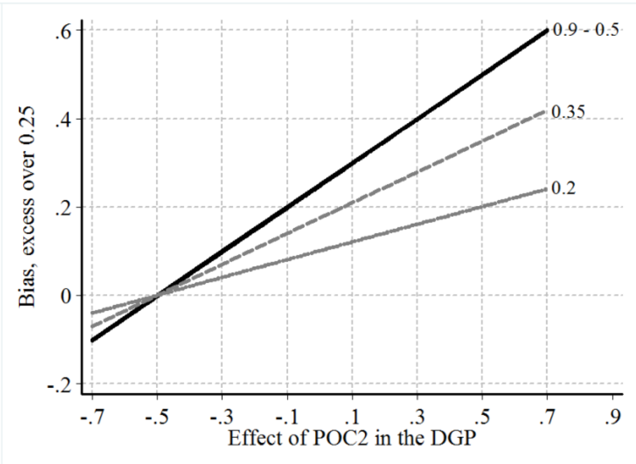


Table 4. Experiments with q and γ_2 – the main range

For	POC2	Mean	Right augmentation			Wrong augmentation			PET meta average		
q	γ_2	$\underline{b}_a$	$\underline{\beta}_{AR}$	*Avs*	*Fs*	$\underline{\beta}_{AW}$	*Avs*	*Fs*	$\underline{\beta}_M$	*Avs*	*Fs*
0.9	-0.7	0.160	**0.249**	96	96	0.148	0	96	0.154	0	96
	-0.5	0.250	**0.249**	96	95	**0.248**	96	95	**0.248**	96	95
	-0.3	0.340	**0.249**	96	93	0.348	0	93	0.351	0	84
	0	0.475	**0.249**	96	93	0.499	0	93	0.511	0	31
	0.3	0.610	**0.250**	97	94	0.649	0	94	0.661	0	27
	0.5	0.700	**0.250**	97	92	0.749	0	92	0.747	0	59
	0.7	0.789	**0.250**	97	90	0.849	0	90	0.813	0	91
0.8	-0.7	0.171	**0.248**	96	97	0.148	0	97	0.160	0	95
	-0.5	0.250	**0.248**	96	98	**0.248**	95	98	**0.248**	95	98
	-0.3	0.330	**0.248**	96	96	0.348	0	96	0.352	0	68
	0	0.450	**0.249**	95	95	0.498	0	95	0.517	0	2
	0.3	0.569	**0.249**	95	93	0.648	0	93	0.665	0	1
	0.5	0.649	**0.249**	96	93	0.749	0	93	0.737	0	23
	0.7	0.729	**0.249**	97	93	0.849	0	93	0.774	0	87
0.65	-0.7	0.186	**0.248**	94	96	0.148	0	96	0.173	0	92
	-0.5	0.250	**0.248**	94	96	**0.248**	94	96	**0.248**	93	96
	-0.3	0.315	**0.248**	95	96	0.348	0	96	0.349	0	32
	0	0.412	**0.249**	96	95	0.498	0	95	0.513	0	0
	0.3	0.508	**0.249**	97	95	0.649	0	95	0.651	0	0
	0.5	0.573	**0.249**	97	97	0.749	0	97	0.702	0	5
	0.7	0.637	**0.249**	97	97	0.849	0	97	0.702	0	80
0.5	-0.7	0.200	**0.250**	92	94	0.149	0	94	0.189	1	93
	-0.5	0.250	**0.250**	93	95	**0.250**	95	95	**0.250**	93	96
	-0.3	0.301	**0.250**	91	95	0.350	0	95	0.339	0	22
	0	0.376	**0.250**	91	95	0.500	0	95	0.491	0	0
	0.3	0.451	**0.250**	95	96	0.650	0	96	0.610	0	0
	0.5	0.501	**0.250**	95	95	0.750	0	95	0.640	0	7
	0.7	0.551	**0.250**	94	96	0.850	0	96	0.612	0	79
0.35	-0.7	0.215	**0.250**	97	95	0.180	0	95	**0.240**	90	57
	-0.5	0.250	**0.250**	95	94	**0.250**	91	94	**0.250**	96	93
	-0.3	0.285	**0.250**	95	93	0.320	0	93	0.323	2	51
	0	0.338	**0.250**	94	94	0.425	0	94	0.452	0	2
	0.3	0.390	**0.250**	95	93	0.529	0	93	0.546	0	4
	0.5	0.425	**0.250**	95	94	0.599	0	94	0.560	1	28
	0.7	0.460	**0.250**	95	96	0.669	0	96	0.520	41	85
0.2	-0.7	0.230	**0.250**	92	92	0.210	7	92	**0.253**	93	46
	-0.5	0.250	**0.250**	92	92	**0.250**	93	92	**0.250**	91	92
	-0.3	0.270	**0.250**	91	91	0.291	4	91	0.299	86	85
	0	0.301	**0.250**	94	92	0.351	0	92	0.388	23	76
	0.3	0.331	**0.250**	91	92	0.411	0	92	0.447	33	26
	0.5	0.351	**0.250**	92	93	0.451	0	93	0.447	83	0
	0.7	0.371	**0.250**	91	92	0.491	0	92	0.411	84	0

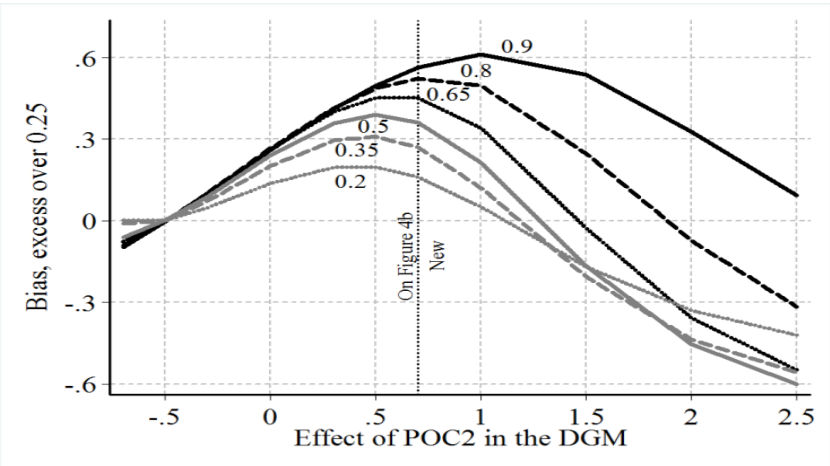
Note: DGP = EM = 3, β = 0.25, N = 100 and $\rho = \gamma_1$ = 0.5

Table 4b. Experiments with q and γ_2 – extended range

For	POC2	Mean	Right augmentation			Wrong augmentation			PET meta average		
q	γ_2	$\underline{b_a}$	$\underline{\beta_{AR}}$	*Avs*	*Fs*	$\underline{\beta_{AW}}$	*Avs*	*Fs*	$\underline{\beta_M}$	*Avs*	*Fs*
0.9	1.0	0.924	0.250	96	91	1.000	0	91	0.861	0	68
	1.5	1.148	0.251	97	93	1.250	0	93	0.787	0	0
	2.0	1.373	0.251	97	92	1.501	0	92	0.578	2	0
	2.5	1.597	0.251	97	93	1.751	0	93	0.344	66	0
0.8	1.0	0.849	0.250	99	93	0.999	0	93	0.746	0	53
	1.5	1.048	0.250	99	91	1.249	0	91	0.497	5	0
	2.0	1.248	0.250	98	92	1.500	0	92	0.176	70	0
	2.5	1.447	0.250	98	92	1.750	0	92	-0.067	1	0
0.65	1.0	0.735	0.250	98	98	0.999	0	98	0.591	0	34
	1.5	0.897	0.250	98	98	1.249	0	98	0.222	94	0
	2.0	1.059	0.250	99	99	1.499	0	99	-0.108	0	0
	2.5	1.220	0.250	99	99	1.749	0	99	-0.299	0	0
0.5	1.0	0.626	0.250	93	96	1.001	0	96	0.463	0	28
	1.5	0.751	0.251	92	95	1.251	0	95	0.082	18	0
	2.0	0.876	0.251	92	95	1.502	0	95	-0.203	0	0
	2.5	1.002	0.251	93	96	1.752	0	96	-0.351	0	0
0.35	1.0	0.515	0.250	93	95	1.000	0	95	0.371	41	27
	1.5	0.603	0.251	93	94	1.251	0	94	0.044	2	0
	2.0	0.691	0.251	93	93	1.501	0	93	-0.187	0	0
	2.5	0.780	0.251	93	93	1.752	0	93	-0.307	0	0
0.2	1.0	0.402	0.250	94	96	1.001	0	96	0.301	81	47
	1.5	0.452	0.250	94	95	1.251	0	95	0.080	2	0
	2.0	0.502	0.250	94	95	1.502	0	95	-0.080	0	0
	2.5	0.553	0.250	94	95	1.752	0	95	-0.171	0	0

Note: DGP = EM = 3, β = 0.25, N = 100 and $\rho = \gamma_1$ = 0.5

Figure 4e. Extended version of Figure 4b



As in Section 3 the range for the POC is extended to study the paths to the β_M-lines, but as the paths seems simpler the extension is less dense. The extended version of figure 4b is as shown.

5. Combinations of DGPs and EMs: First set of experiments

Table 5 show one set of simulations with all combinations of equations (2a), (2b) and (3) from Table 1. The parameter set is: $\gamma_1 = 0.75$, $\gamma_2 = -0.5$, $\rho = 0.7$ and $q = 0.5$. Section 7 looks at a different set of parameters.

The key point to note is that only four of the 9 *x* 4 = 36 averages are right. This is we get it right in 11 % of the estimates. It only happens when the EM is equal to or contains the DGP, and in these cases the right augmentation is needed.

The 9 lines are all illustrated by one funnel. The funnels are rather typical of the 100 calculated for the Table. Line (1) in the table is thus illustrated on Figure 5a. It is a case where the DGP = EM so that the right augmentation finds the true value of β. The funnel consists of two funnels, the true funnel is at the left hand and it is the 'lowest'. The right hand funnel is false and actually higher. This gives a problem. The PET stars in the average for low precision, which is at the bottom of the left hand (right) funnel and due to the curvature the PET overshoots the false funnel and becomes a little worse than the wrongly augmented meta average.

The second case in Line (2) of the table and Figure 5b is more interesting. Here the EM is wrong in the sense that the true POC does not enter. Here all four averages are rather similar and all rather bad. Here the funnel looks ideal, all FAT are satisfactory. Clearly the meta analysis does not discover that something is amiss.

Table 5. All combinations of DGP and EM's, with one set of parameters

Row	DGP	EM	Mean	Right augmentation			Wrong augmentation			PET meta average		
			$\underline{b}_a$	$\underline{\beta}_{AR}$	*Avs*	*Fs*	$\underline{\beta}_{AW}$	*Avs*	*Fs*	$\underline{\beta}_M$	*Avs*	*Fs*
(1)	2a	2a	0.514	**0.249**	95	96	0.775	0	96	0.804	0	0
(2)	2a	2b	0.775	0.774	0	97	0.774	0	97	0.775	0	96
(3)	2a	3	0.514	**0.250**	95	95	0.775	0	95	0.949	0	0
(4)	2b	2a	-0.100	-0.101	0	94	-0.101	0	94	-0.101	0	96
(5)	2b	2b	0.075	**0.249**	92	94	-0.101	0	94	-0.163	0	0
(6)	2b	3	0.075	**0.250**	93	95	-0.101	0	95	-0.231	0	0
(7)	3	2a	0.164	-0.101	0	95	0.424	0	95	0.467	0	0
(8)	3	2b	0.600	0.774	0	97	0.424	0	97	0.355	0	0
(9)	3	3	0.338	**0.250**	94	95	0.425	0	95	0.475	0	0
Means			0.328	0.288	52.1	95.3	0.366	0	95.3	0.370	0	21.3

Parameters: $\gamma_1 = 0.75$, $\gamma_2 = -0.5$, $\rho = 0.7$ and $q = 0.5$. Entries are averages of 100 funnels of 500 estimates.

Figure 5a: Row 1 from Table 5

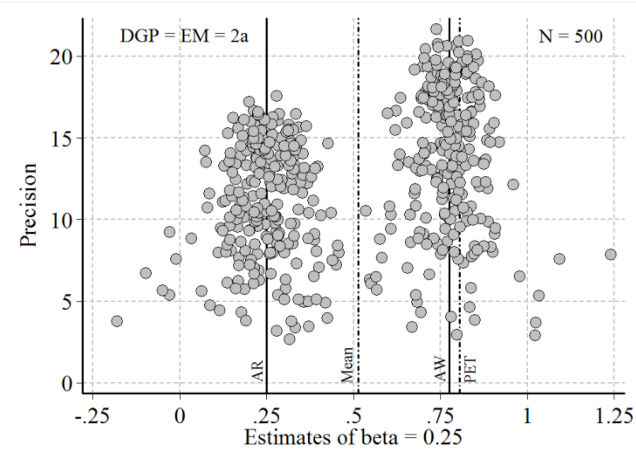


Figure 5b: Row 2 from Table 5

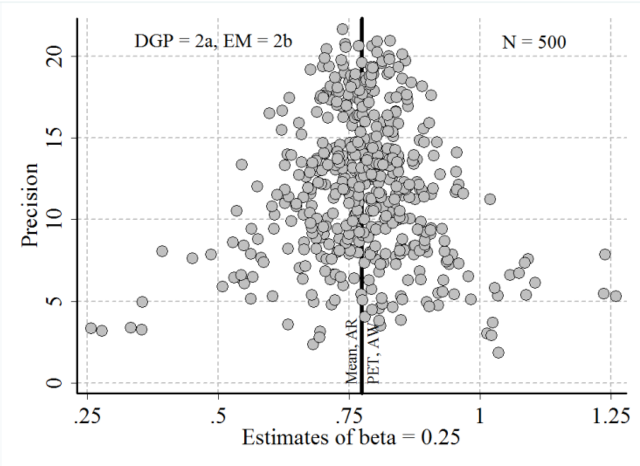


Figure 5c: Row 3 from Table 5

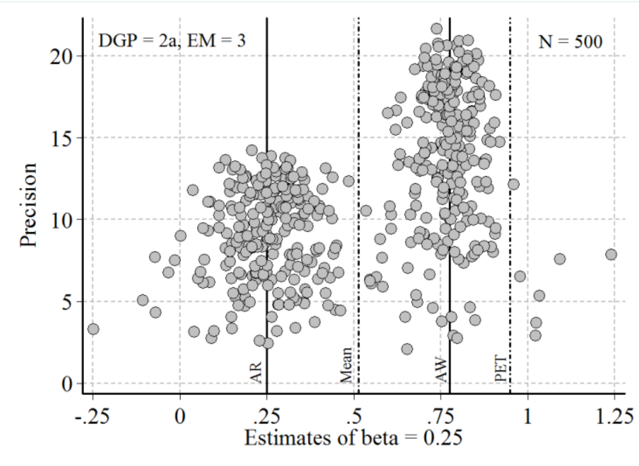


Figure 5d: Row 4 from Table 5

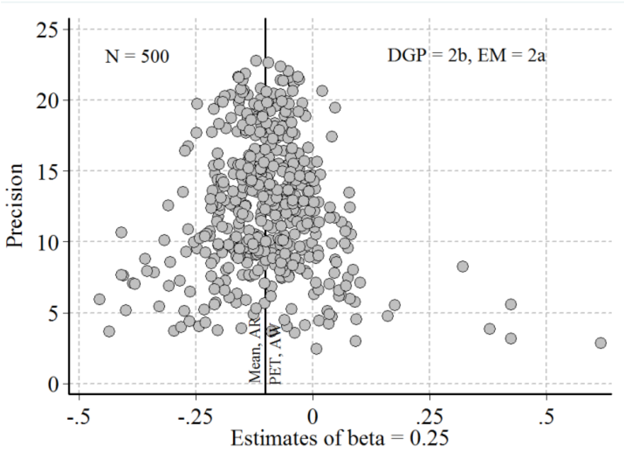


Figure 5e: Row 5 from Table 5

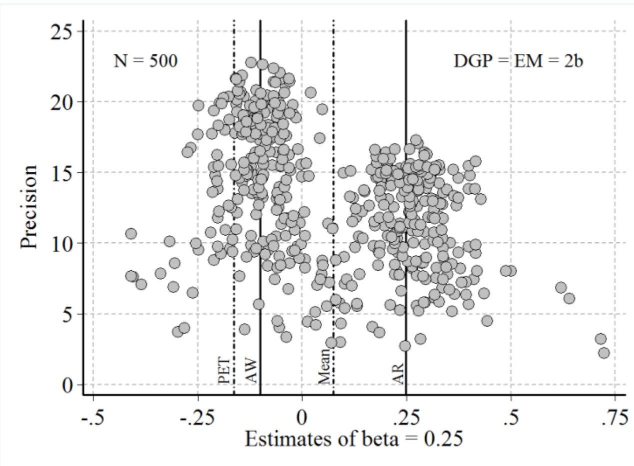


Figure 5f: Row 6 from Table 5

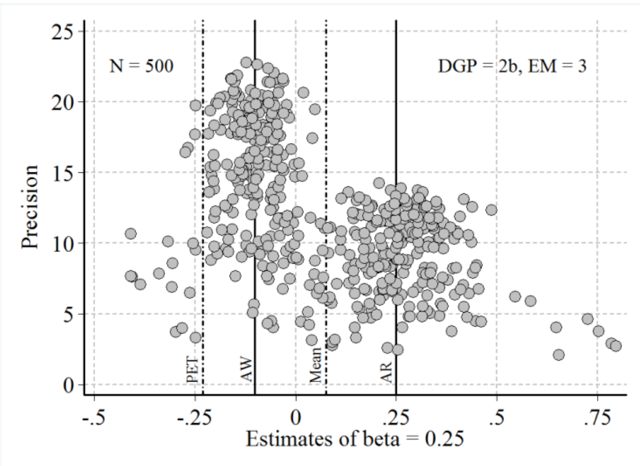


Figure 5g: Row 7 from Table 5

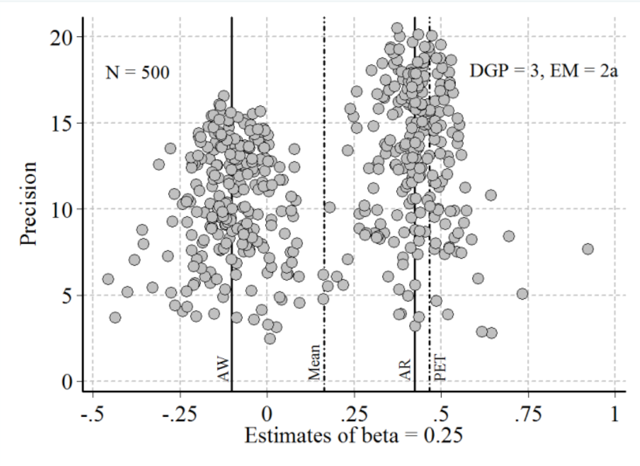


Figure 5h: Row 8 from Table 5

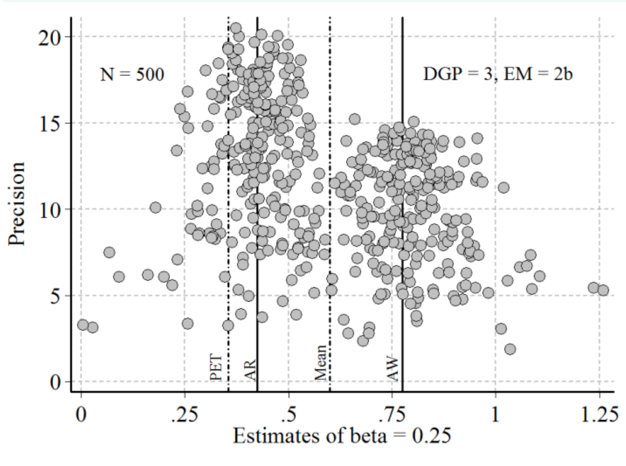
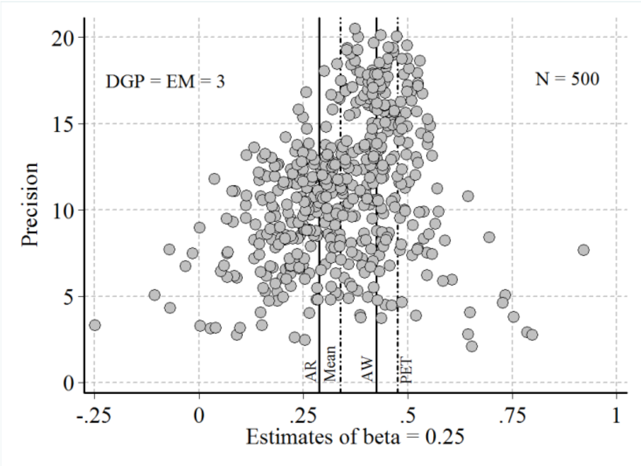


Figure 5i: Row 9 from Table 5



The story from line (1) in repeated, but at a lower level of significance in line (3) as illustrated on Figure 5c. Here the irrelevant POC2 increases extra noise, but otherwise the results are similar. The next three lines are the symmetric case where the DGB is equation 2b instead of 2a.

Lines (7) and (8) are illustrated in Figures 5g and 5h that look deceptively like the two previous two-topped cases. Further, it looks as one of the two peaks may be the right one, but as one POC is missing both peaks are wrong.

Finally, in the confusing case of Figure 5h it looks as if there is only one peak but clearly it is asymmetric. Still the right augmentations works and the PET fails.

6 Censoring some of the cases from section 4

In this section the cases from lines (1), (2) (6) and (9) in Table 5 are censored. This is the complex case of a few POC biases and censoring combined. This gives rather confusing results.

Case (1) from Table 5 is now in lines (1) and (2). Line (1) censors about half of the true funnel and line (2) censors the false funnel fully. The results are as should be predicted, censoring makes the results worse.

In the two cases of line (2) from Table 5 the PET behaves exactly as it should. But as a POC is missing the results are not right. However, the adjustment is right. In the case from line (6) in Table 5, censoring makes the false peak go away, so everything works nicely.

Table 6. Censored versions of some of the cases from Table 5

Row	DGP. EM	Cen-Sored	Mean $\underline{b_a}$	Right augmentation $\underline{\beta_{AR}}$	*Avs*	*Fs*	Wrong augmentation $\underline{\beta_{AW}}$	*Avs*	*Fs*	PET meta average $\underline{\beta_M}$	*Avs*	*Fs*
1a	2a, 2a	0.25	0.313	**0.249**	91	90	0.500	0	90	0.345	0	73
1b	2a, 2a	0.5	0.501	0.499	0	96	0.500	0	96	0.499	0	96
2a	2a, 2b	0.6	0.313	**0.249**	91	94	0.500	0	94	0.413	0	0
2b	2a, 2b	0.7	0.500	0.500	0	89	0.500	0	89	0.500	0	89
6a	2b, 3	0	0.313	**0.249**	94	95	0.500	0	95	0.346	0	75
6b	2b, 3	0.2	0.313	**0.249**	92	95	0.500	0	95	0.413	0	0
9a	3, 3	0.2	0.563	0.500	0	90	0.751	0	90	0.603	0	58
9b	3, 3	0.3	0.564	0.499	0	97	0.750	0	97	0.603	0	64
9c	3, 3	0.4	0.376	**0.249**	92	94	0.750	0	94	0.489	0	16
Averages			0.360	0.360	51.1	93.3	0.583	0.0	93.3	0.468	0	21.3

Note: Row 1a and 1b are two censored version of Row 1 from Table 5. Parameters: γ_1 = 0.75, γ_1 = -0.5, ρ = 0.7 and q = 0.5. Entries are averages of 100 funnels of estimates. The number of regressions R made per funnel is assessed to give 500 observations after the censoring.

Figure 6a: Row 1a from Table 6

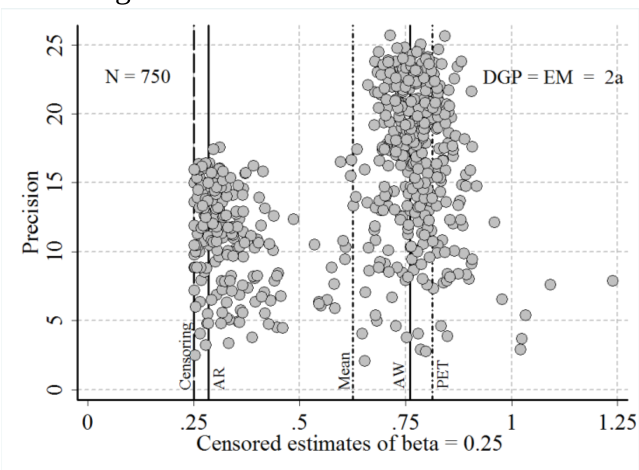


Figure 6b: Row 1a from Table 6

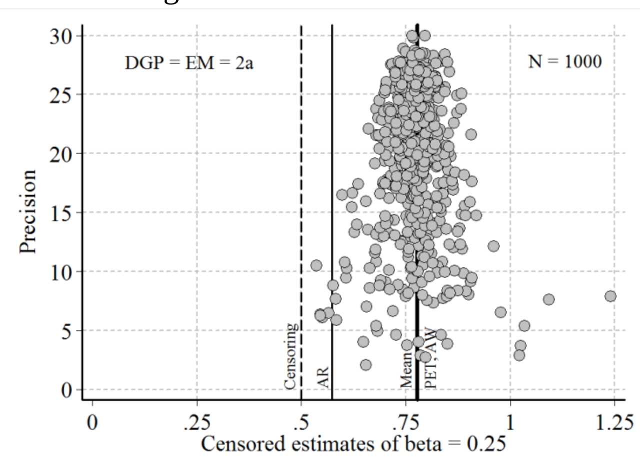


Figure 6c: Row 2a from Table 6

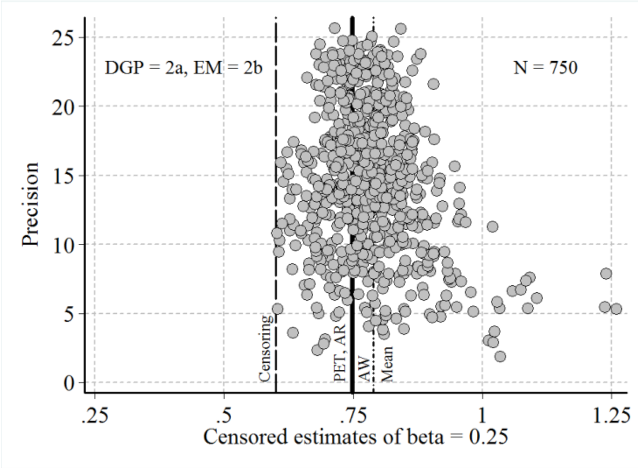


Figure 6d: Row 2b from Table 6

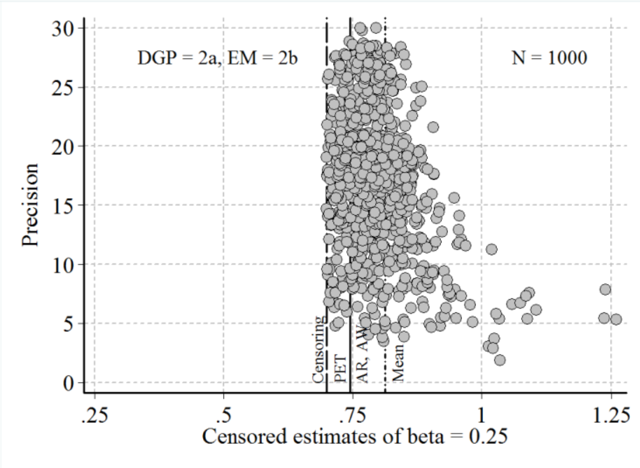


Figure 6e: Row 6a from Table 6

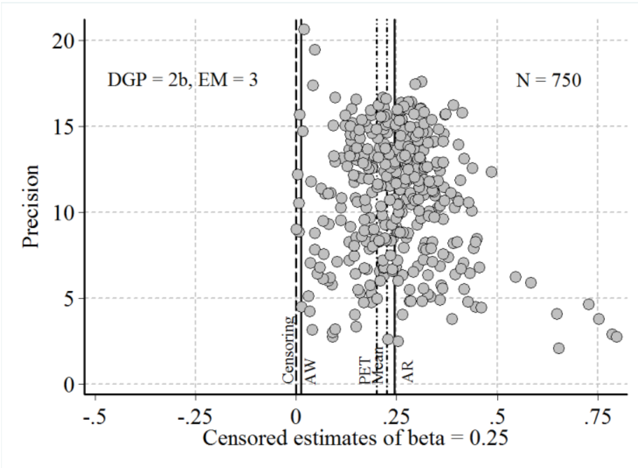


Figure 6f: Row 5b from Table 6

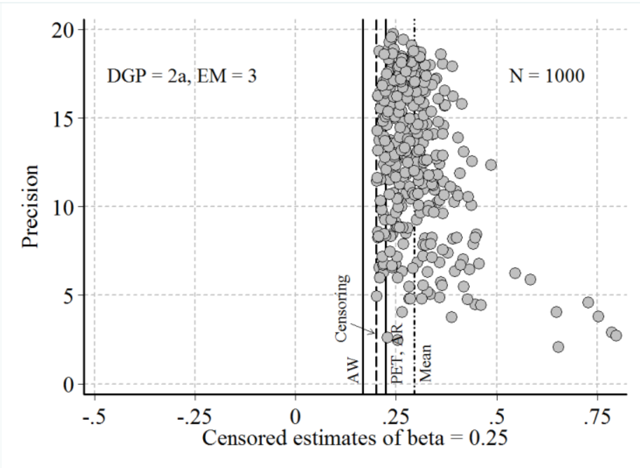


Figure 6g: Row 9a from Table 6

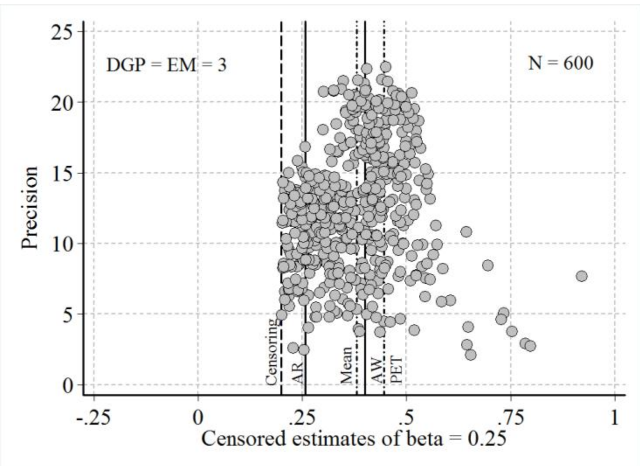


Figure 6h: Row 9b from Table 6

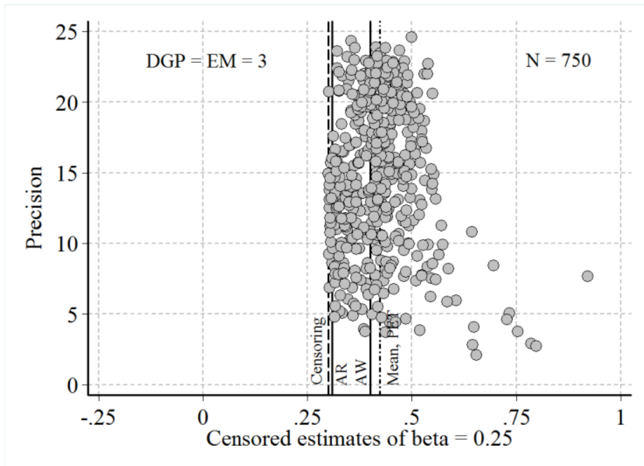
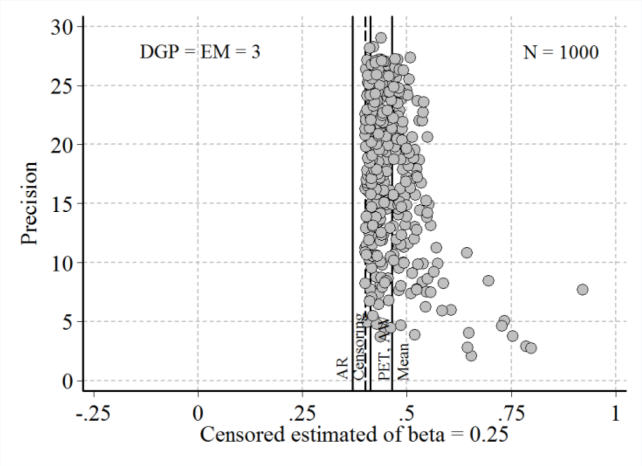


Figure 6i: Row 9c from Table 6



Finally, in the case from line (9) of Table 5 the results are rather confusing. Here censoring is not rightly detected by the PET. However the PET moves toward the mean.

7. Combinations of DGPs and EMs: Second set of experiments

This section makes the same analysis as in section 5 but for another set of parameters γ_1 = 0.5, γ_1 = 0.5, ρ = 0.5 and q = 0.25. I have chosen parameters that in Sections 3 and 4 give results that differ substantially. However, the pattern in the results of the present exercise is not very different from the one in section 5.

Table 7. All combinations of DGP and EM's, with one set of parameters

Row	DGP	EM	Mean	Right augmentation			Wrong augmentation			PET meta average		
			$\underline{b}_a$	$\underline{\beta}_{AW}$	*Avs*	*Fs*	$\underline{\beta}_{AR}$	*Avs*	*Fs*	$\underline{\beta}_{M}$	*Avs*	*Fs*
1	2a	2a	0.313	**0.249**	91	90	0.500	0	90	0.345	0	73
2	2a	2b	0.501	0.499	0	96	0.500	0	96	0.499	0	96
3	2a	3	0.313	**0.249**	91	94	0.500	0	94	0.413	0	0
4	2b	2a	0.600	0.600	0	89	0.601	0	89	0.600	0	90
5	2b	2b	0.338	**0.249**	95	95	0.600	0	95	0.327	13	97
6	2b	3	0.338	**0.250**	90	94	0.600	0	94	0.419	0	20
7	3	2a	0.663	0.600	0	91	0.851	0	91	0.709	0	48
8	3	2b	0.589	0.499	0	96	0.850	0	96	0.594	0	97
9	3	3	0.402	**0.249**	92	94	0.850	0	94	0.449	0	85
Means			0.451	0.383	51.0	93.2	0.650	0.0	93.2	0.484	1.4	67.3

Parameters: γ_1 = 0.5, γ_1 = 0.5, ρ = 0.5 and q = 0.25. Entries are averages of 100 funnels of 500 estimates.

Figure 7a: Row 1 from Table 7

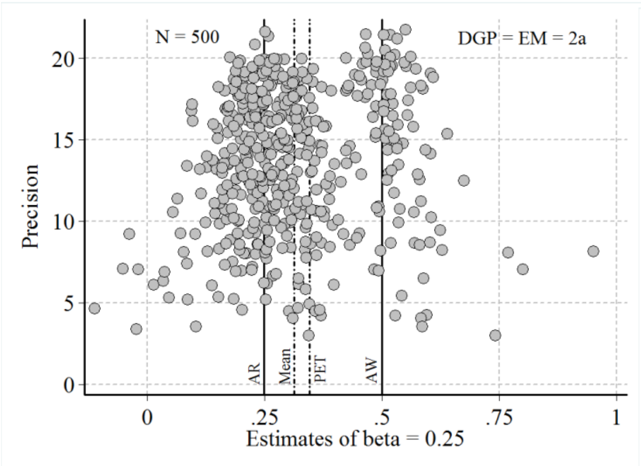


Figure 7b: Row 2 from Table 7

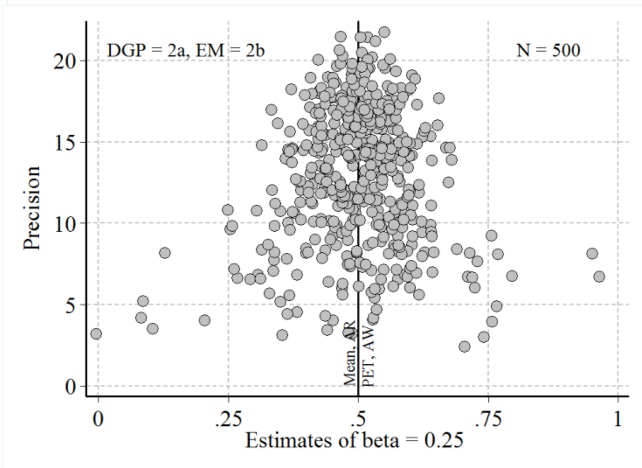


Figure 7c: Row 3 from Table 7

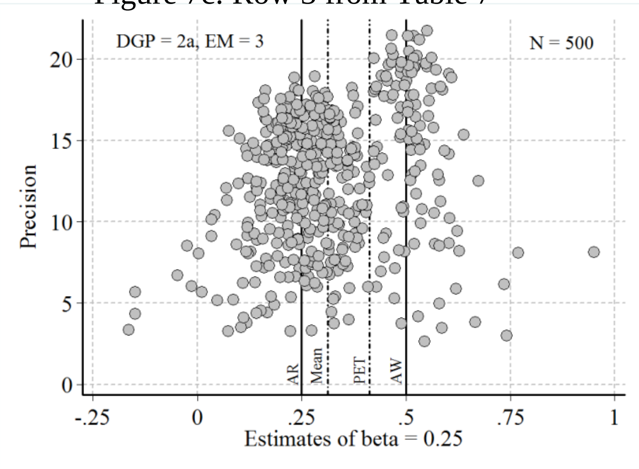


Figure 7d: Row 4 from Table 7

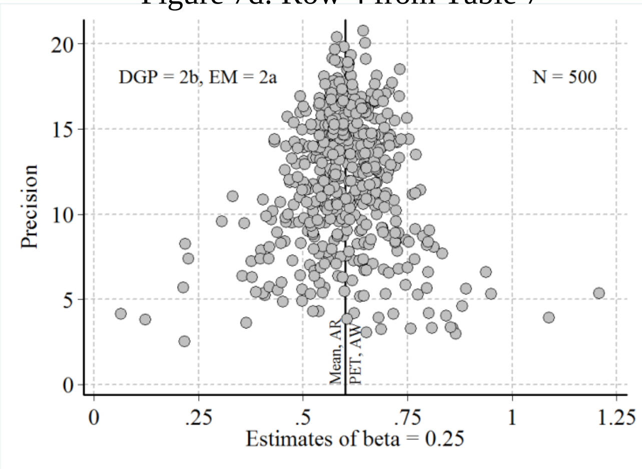


Figure 7e: Row 5 from Table 7

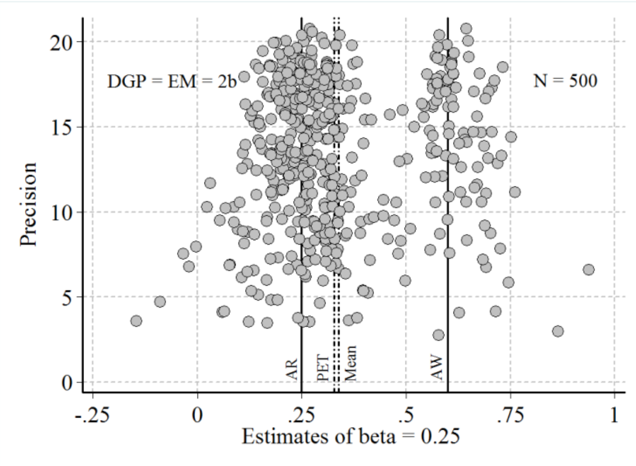


Figure 7f: Row 6 from Table 7

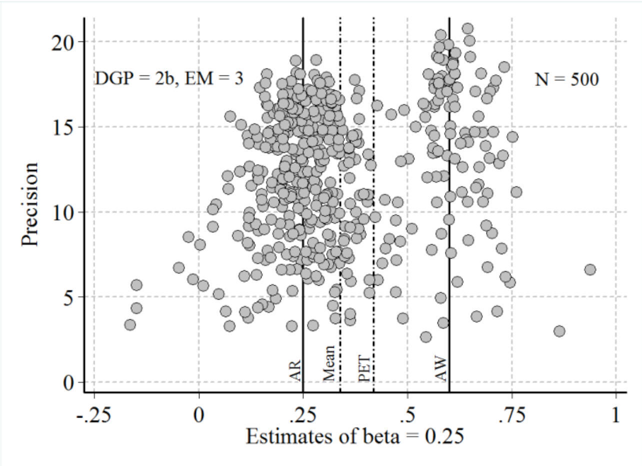


Figure 7g: Row 7 from Table 7

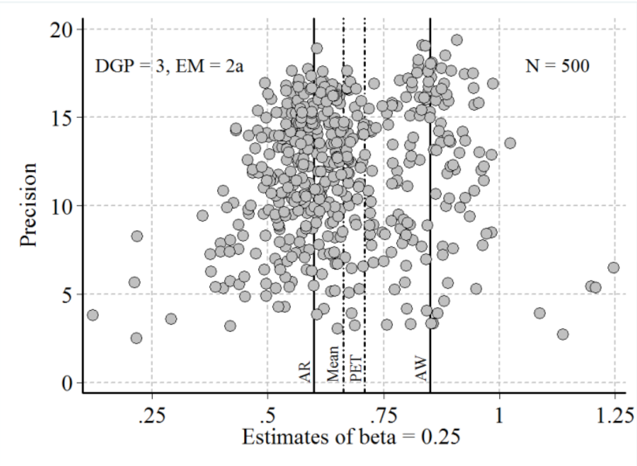


Figure 7h: Row 8 from Table 7

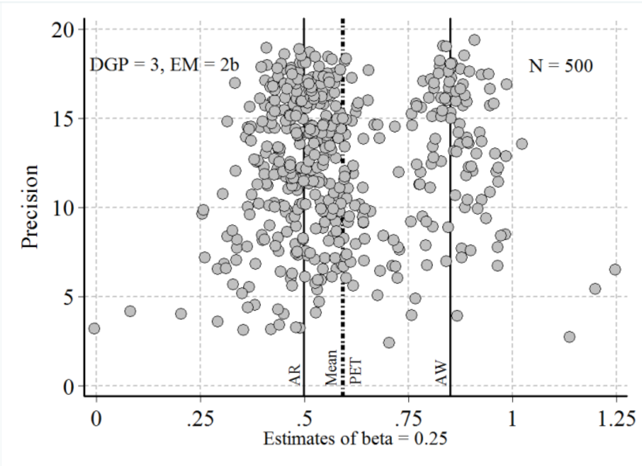
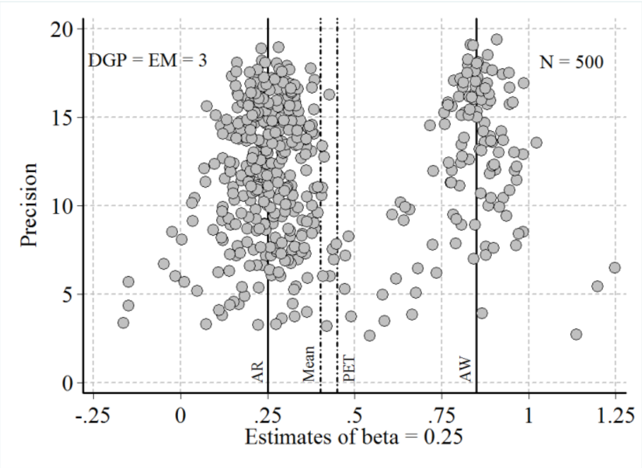


Figure 7i: Row 9 from Table 7



8. The simulation program

The simulation program is written as three stata do-programmes. They run three loops as covered in Table 8a.

Table 8a. The three loops of the simulation programs

Outer loop: ***Set R*** to generate R funnels: [stata: **running**, set N and R, and all parameters

Middle loop: Set N to generate one funnel with N estimates using sample length M = 20, 21, …, 19 + N simulated observations. [stata: **calibrate**, set N, and all parameters]

Inner Loop: One primary regression to get ***results i***. Using M observations to

Generate series x, z_1 and z_2. All have zero mean and a standard deviation, sd, ***set***. Also, z_1 and z_2 are correlated with x. The correlation is the same, it is ***set***.

DGP, calculate y by one of 5 possible equation from x and z_1 or z_2. The equation is ***set***. In all experiments $\beta = 0.25$ is the effect of interest.

EM, estimate $b \approx \beta$. By an equation that is ***set*** from the same 5 possibilities as the DGP. The POC (z_1 or z_2) is included with the chance ω, ***set***.

Keep: Results $i = (b_i, s_i, \omega_i, t_i, p_i, \varphi_i)$. Altogether 6 variables are kept.

[stata: **programs_meta**, called by the other two programs, all parameters should be set]

Back to middle loop: After N runs of inner loop we have the **(N x 6)** results matrix, which is the information for one funnel. Estimate the average, $\underline{b}$, FAT-PET MRA giving (β_M , β_F) and two augmentations, using φ, giving (β_{AR}, β_{FR})and ω giving (β_{AW}, β_{FW}) and six counts; $Avs1$ counts if $\beta_M \approx \beta$, $Fs1$ count if $\beta_F \approx 0$, $Avs2$ count if $\beta_{AR} \approx \beta$, $Fs2$ count if $\beta_{FR} \approx 0$, $Avs3$ count if $\beta_{AW} \approx \beta$ and finally $Fs3$ count if $\beta_{FW} \approx 0$. Where all the "≈" means that it is not rejected that the said equation holds at the 5 % level of significance.

Keep: the estimates of $\underline{b}$, β_M, β_{AR}, β_{AW}, and the latest values of $Avs1$, $Fs1$, $Avs2$, $Fs2$, $Avs3$, $Fs3$.

Back to outer loop: After R runs of middle loop we have the (***R*** *x* **4**) result matrix of the $\underline{b}$, β_M, β_{AR}, β_{AW}, and the six counts: $Avs1$, $Fs1$, $Avs2$, $Fs2$, $Avs3$, $Fs3$. Calculate the average of the four columns. Report the 10 statisics.

The program consists of three do-files:

programs_meta, which runs the middle and inner level. This program is not run separately, but only by the other two programs. It uses the simulation programs of stata.

calibration runs the middle and inner level to generate one funnel. It calls programs_meta and generate a funnel to help to calibrate the big running. All funnels shown are from the calibration do-file. 11 parameters have to be set see Table 8b.

running runs all three levels. It calls programs_meta and runs R funnels. 12 parameters have to be set here see Table 8b.

Table 8b show what the parameters that have to be set are and where they have to be set. Also, it lists the values tried in the experiments.

Table 8b. Parameters to be set

To be set	The content of the parameter to be set	Set in calibration	Set in running	Normally set at	Experiments in Section	Experiments
R	Number of funnels		X	100 or 1000		
N	Number of estimates for funnel	X	X	500		
β	Effect if interest, coefficient on x	X	X	0.25		
Sd ε	Std of *x* in DGP	X	X	1	2	0.5, 2, 3, 4, 6, 7
Sd z_1	Std of POC 1 (z_1) in DGP	X	X	1		
Sd z_2	Std of POC 2 (z_2) in DGP	X	X	1		
γ_1	Coefficient on POC 1 in DGP	X	X	0.5 or 0.75	3	Big range
γ_2	Coefficient on POC 2 in DGP	X	X	-0.5 or 0.5	4	-0.5, -0,25, 0, … , 1
ρ	Correlation of z1 and z2 with x in DGP	X	X	0.5 or 0.7	3	0.9 to 0
q	Probability of inclusion of POCs in EM	X	X	0.5	4	0.9 t0 0.2
DGP	Data generating process: (1), (2a) (2b) or (3)	X	X	Both (1) or (3)	5,6,7	All combinations of (2a) (2b) and (3)
EM	Estimation model, using OLS	X	X			

To run *calibration* takes about 20 seconds for a funnel with N = 500 points.

To run *running* with 100 similar funnels takes about 15 minutes

Conclusions

This appendix reports 182 experiments. For each one 50,000 simulated regressions have been made. This gives a total of 9,100,000 regressions, and as some experiments have needed more due to censoring and the 40 funnels are independently generated it adds to more than 9½ million regressions.

I take the case with zero POCs to represent the case with many POCs. And on the other end I have covered the case with 1 and with 2 POCs. Since these cases give rather simple patterns in the results when to the parameter variation tried, I assess that we understand these cases rather well. But it will, of course, be easy to suggest other combinations of the parameters, e.g., by untying ρ and q for the two POCs.

However, what will really blow up the number of experiments is to consider cases with 3, 4 or even more POCs.