PNAS Nexus, 2026, **5**, pgag150
https://doi.org/10.1093/pnasnexus/pgag150
Published: 5 May 2026
Brief report

No evidence that nonincentivized behavioral interventions effectively mitigate climate change after adjusting for publication bias

Adam Hardaker[1], Igor Asanov[1], František Bartoš[2], and Stephan B. Bruns[1,3,4]*

[1]Department of Economics, International Center for Higher Education Research (INCHER), University of Kassel, Kassel 34125, Germany
[2]Faculty of Social Sciences, Institute of Economic Studies, Charles University, Prague 110 01, Czech Republic
[3]Centre for Environmental Sciences, Hasselt University, Hasselt 3500, Belgium
[4]Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA 94305, USA
*To whom correspondence should be addressed: Email: stephan.bruns@uhasselt.be
Edited By Valerio Capraro

Abstract

Behavioral interventions on citizens are often promoted as a low-cost route to induce environmentally friendly behavior, yet published estimates of their effectiveness are highly variable and prone to selective reporting. We reanalyzed the evidence of nonincentivized behavioral interventions on citizens. We applied robust Bayesian meta-analysis (RoBMA), averaging across a full set of publication bias–adjusted models, to the 144 effect estimates (91 studies) compiled by Nisa et al. (2019). After accounting for publication bias and model uncertainty using multilevel RoBMA, the data strongly favor a zero average effect. The posterior probability that the meta-analytic mean equals zero is 0.984, and the Bayes factor comparing a zero mean to a nonzero mean is $BF_{01} = 63.5$. Accordingly, the previously reported mean benefit of behavioral interventions on households and individuals may largely reflect publication bias and potentially other small-study effects. There is evidence for small between-study heterogeneity, indicating that some specific interventions might have an effect. These results suggest that, on average, behavioral interventions without incentives on households and individuals are unlikely to deliver material climate benefits.

Keywords behavioral interventions, climate change mitigation, publication bias, robust Bayesian meta-analysis, proenvironmental behavior

Introduction

Citizens' consumption choices contribute a sizable share of global greenhouse gas emissions, so even small behavioral shifts can aid climate goals (1). Behavioral interventions, such as freedom-preserving choice architecture changes, are attractive because they appear inexpensive and politically acceptable (2). Reported effects, however, are heterogeneous and based almost entirely on published studies that are potentially subject to publication bias[a] which may distort published evidence, particularly in situations where the true effect is zero or negligibly small (3).

A comprehensive compilation of the effects of nonincentivized behavioral interventions on households and individuals to mitigate climate action, produced in a meta-analysis by Nisa et al. (4), contains 144 effect estimates drawn from 91 field studies, most of them randomized controlled trials. Their conventional random-effects model without any correction for publication bias suggested a mean standardized impact of $d = -0.093$ (i.e. a 54% probability that a randomly chosen treated observation is below the mean of the control group), yet a sensitivity analysis shows that the true effect size may be as low as $d = -0.028$ (51%). Other researchers remeasure the effect of behavioral interventions with the same data and report an effect even twice as large, $d = -0.204$ (58%) (5). These divergent figures highlight sensitivity to analytic choices, particularly how publication bias is modeled and addressed.

Rather than selecting a single "best" specification, we evaluate robustness across a set of analytic alternatives. This is done by implementing a recently developed method of robust Bayesian meta-analysis (RoBMA) (6), which averages across an ensemble of bias-correction models, weighting each model by its predictive performance. In the multilevel version of RoBMA used here, estimates are clustered by study to account for dependence among multiple effect sizes per study. This approach yields posterior

Received: August 25, 2025. **Accepted:** April 19, 2026

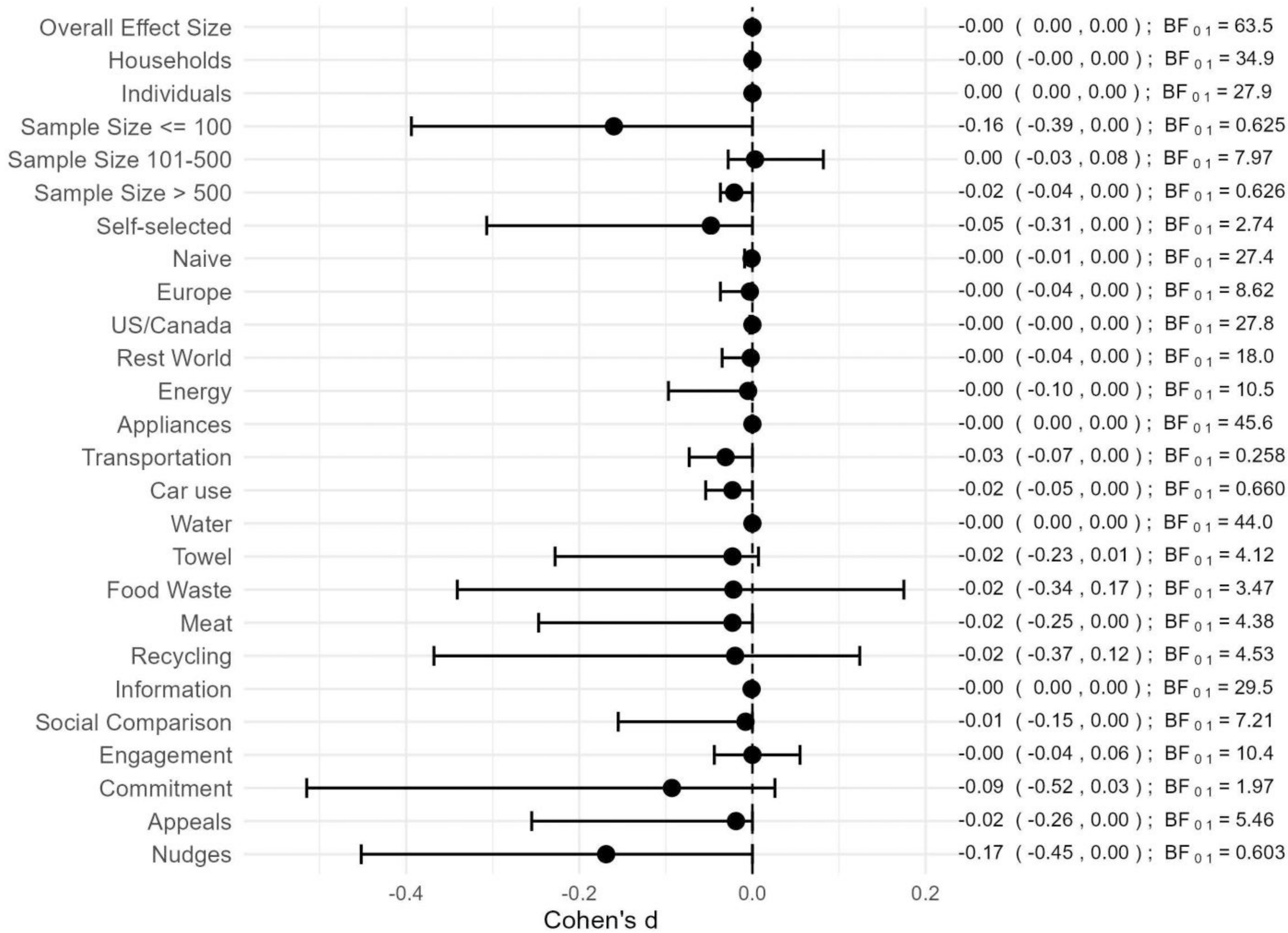

Figure 1 Forest plot for the publication bias–adjusted mean effect size from multilevel RoBMA estimates. Publication bias–adjusted mean effect size estimates using the $\text{RoBMA}_{\text{PMSA}}$ model–averaged posterior mean is presented with 95% credible intervals and inclusion Bayes factors (BF), both for the overall sample and across various subgroups. SEs are clustered by study via multilevel RoBMA. The BF_{01} quantifies support for the null hypothesis, with values >1 indicating evidence in favor of the null, while values <1 support the alternative hypothesis (BF_{10} can be derived as $1/BF_{01}$). As a general guideline, BFs between 3 and 10 suggest moderate evidence, while values exceeding 10 provide strong evidence.

inferences that incorporate model uncertainty and are less sensitive to researcher discretion. Using RoBMA on the full dataset, we provide a publication bias–adjusted, policy-relevant estimate of how much nonincentivized behavioral interventions can realistically induce proenvironmental behaviors.

Results

Figure 1 displays the multilevel RoBMA publication bias–adjusted forest plot for the full data and various subgroups as reported in Nisa et al. (2019). The evidence strongly favors a zero average effect. Across all 144 estimates, the posterior probability that the meta-analytic mean equals zero in the multilevel RoBMA is 0.984. The Bayes factor comparing the data under a zero versus a nonzero mean is $BF_{01} = 63.5$, constituting very strong evidence against an average impact of behavioral interventions on citizens' proenvironmental behaviors. However, the model finds decisive evidence for between-study heterogeneity with $\tau = 0.04$ (95% central credible interval [CrI] 0.02, 0.11). While small in magnitude, this indicates that the true effects may vary slightly across studies (i.e. some studies may provide true benefits), even though the model-averaged mean is essentially zero.

Table 1 illustrates the difference between unadjusted and publication bias–adjusted estimates. Column 1 presents the statistical reproduction of the Nisa et al. paper as a baseline. Like the original published results, the unadjusted overall effect is modestly negative ($d = -0.093$), with a CI that excludes zero. Column 2 reports multilevel RoBMA estimates that account for dependence, with $BF_{01} = 63.5$ providing strong evidence for a null effect. Statistically significant negative effects in the random-effects analysis are attenuated once estimates are adjusted for publication bias using RoBMA; this pattern largely extends to the subgroup analysis. For the household and individual subgroups, the model-averaged estimates are essentially zero as well, with Bayes factors favoring the absence of an average effect. Some of the subgroup analyses show Bayes factors in favor of an effect (without correction for multiple testing). Out of the 25 subgroup analyses, this is true for the smallest and largest sample size groups, transportation, including car use and nudges, but with considerably smaller effect sizes compared with the original estimates. Such discrepancies between unadjusted and publication

Table 1 Unadjusted and publication bias–adjusted meta-analytic estimates of nonincentivized behavioral interventions on proenvironmental behavior (Cohen's *d*).

		(1) Random effects (statistical reproduction)			(2) RoBMA$_{PMSA}$ multilevel			(3) RoBMA$_{PMSA}$ multilevel (both PET and PEESE excluded)		
		k	d	CI	μ	CrI	BF$_{01}$	μ	CrI	BF$_{01}$
Overall effect size		144	−0.093	(−0.123, −0.063)	0.000	(0.000, 0.000)	63.5	−0.064	(−0.177, 0.000)	0.776
Sensitivity analysis										
Sample type	Households	66	−0.112	(−0.162, −0.062)	0.000	(−0.003, 0.000)	34.9	−0.008	(−0.147, 0.020)	6.78
	Individuals	78	−0.118	(−0.166, −0.071)	0.000	(0.000, 0.000)	27.9	−0.008	(−0.063, 0.000)	4.42
Sample size per condition	<100	82	−0.335	(−0.411, −0.258)	−0.160	(−0.394, 0.000)	0.625	−0.223	(−0.400, 0.000)	0.214
	100–500	45	−0.137	(−0.194, −0.080)	0.003	(−0.028, 0.082)	7.97	−0.101	(−0.192, 0.000)	0.246
	>500	17	−0.028	(−0.035, −0.021)	−0.021	(−0.037, 0.000)	0.626	−0.013	(−0.037, 0.000)	1.64
Self-selection	Self-selected	79	−0.278	(−0.348, −0.209)	−0.048	(−0.307, 0.000)	2.73	−0.192	(−0.343, 0.000)	0.227
	Naive	65	−0.040	(−0.061, −0.019)	−0.001	(−0.009, 0.000)	27.4	−0.001	(−0.031, 0.000)	20.5
Region	Europe	43	−0.210	(−0.297, −0.123)	−0.003	(−0.037, 0.000)	8.62	−0.022	(−0.175, 0.000)	2.88
	US/Canada	78	−0.108	(−0.155, −0.060)	0.000	(−0.003, 0.000)	27.8	−0.046	(−0.239, 0.000)	2.01
	Rest World	23	−0.059	(−0.092, −0.027)	−0.002	(−0.035, 0.000)	18.0	−0.002	(−0.038, 0.000)	16.2
Behavior										
Energy		47	−0.094	(−0.157, −0.031)	−0.005	(−0.097, 0.000)	10.5	−0.007	(−0.139, 0.002)	8.44
	Appliances	7	−0.008	(−0.041, 0.025)	0.000	(0.000, 0.000)	45.6	0.000	(0.000, 0.000)	47.6
Transportation		21	−0.036	(−0.039, −0.034)	−0.031	(−0.073, 0.000)	0.258	−0.035	(−0.087, 0.000)	0.147
	Car use	17	−0.036	(−0.039, −0.034)	−0.023	(−0.054, 0.000)	0.660	−0.029	(−0.054, 0.000)	0.290
Water		42	−0.052	(−0.085, −0.019)	0.000	(0.000, 0.000)	44.0	−0.002	(−0.022, 0.000)	9.77
	Towel	18	−0.168	(−0.281, −0.054)	−0.023	(−0.228, 0.007)	4.12	−0.047	(−0.261, 0.000)	2.09
Food waste		4	−0.231	(−0.713, 0.251)	−0.022	(−0.341, 0.175)	3.47	−0.037	(−0.361, 0.040)	3.40
Meat		7	−0.239	(−0.513, 0.035)	−0.023	(−0.247, 0.000)	4.39	−0.02	(−0.225, 0.000)	4.81
Recycling		23	−0.457	(−0.626, −0.287)	−0.020	(−0.368, 0.124)	4.53	−0.439	(−0.657, 0.000)	0.036
Intervention										
Information		53	−0.048	(−0.079, −0.017)	−0.001	(0.000, 0.000)	29.5	−0.001	(−0.033, 0.000)	18.4
Social comparison		32	−0.077	(−0.133, −0.022)	−0.008	(−0.155, 0.000)	7.21	−0.021	(−0.211, 0.007)	4.01
Engagement		38	−0.253	(−0.365, −0.140)	0.000	(−0.044, 0.055)	10.4	−0.274	(−0.411, 0.000)	0.029
	Commitment	10	−0.426	(−0.659, −0.193)	−0.093	(−0.515, 0.026)	1.97	−0.213	(−0.583, 0.046)	0.528
Appeals		10	−0.266	(−0.498, −0.034)	−0.019	(−0.255, 0.000)	5.46	−0.039	(−0.320, 0.000)	2.942
Nudges		11	−0.352	(−0.508, −0.196)	−0.169	(−0.452, 0.000)	0.603	−0.186	(−0.449, 0.000)	0.457

The first column (random effects) presents the meta-analytic estimates (Cohen's *d*) with 95% CIs based on clustered SEs. These estimates are not adjusted for publication bias. The second and third columns (RoBMA multilevel and RoBMA multilevel [both PET and PEESE excluded]) provide publication bias–adjusted model-averaged posterior mean effect sizes (with 95% CrI) and corresponding Bayes factors for the absence of the effect (BF$_{01}$), accounting for dependence via study-level clustering. Each subdomain was analyzed separately to allow for a formal test of the presence (or absence) of the effect within that category and to permit publication bias parameters to vary across subdomains. As a note, the publicly available dataset has a different number of studies for "transportation," "car use," "commitment," and "appliances," explaining some of the discrepancies between the statistical reproduction of results and the published original estimates. The overall study count is unchanged.

bias–adjusted estimates highlight how accounting for publication bias can substantially alter conclusions and suggest caution in drawing strong inferences from the unadjusted estimates. Results were robust to both prior choice and dependence: rescaling the default priors (0.5×–2×) kept the overall posterior mean at 0.000 (BF_{01} = 30.2 to 116.6), and a 200-iteration random one-estimate-per-primary-study resampling yielded median μ = −0.001 (2.5th–97.5th: −0.010 to 0.000; BF_{01} median 26.45).

Posterior model weights for the overall analysis concentrate almost entirely on the precision effect test (PET; posterior weight ≈ 1.00). The replication package contains a funnel plot to show that PET is well specified, but we also estimate a multilevel RoBMA exclusively based on selection models (i.e. excluding both PET and precision effect estimate with SE [PEESE]; column 3 in Table 1). This model shows neither evidence for the absence of an average effect nor for its presence (BF_{01} = 0.776), indicating that part of the publication bias adjustment might be driven by small-study effects (8).

Discussion

Taken together, these results indicate that previously reported benefits of nonincentivized behavioral interventions on households and individuals are consistent with substantial publication bias and potentially with other small-study effects. Any real impacts appear context specific and small enough that current evidence lacks the power to distinguish them from zero. Notably, even large sample studies can have limited ability to detect the modest effect sizes typical in behavioral interventions (3). Specifically, random-effects analysis reports an effect size of −0.028 for large sample studies (at least 500 observations per experimental group) and, taking at face value, reliably detecting this effect size with adequate power of 80% requires an effective sample size of at least 40,100 observations (nonconservatively assuming equal allocation fractions, individual level of randomization, 5% significance level, one experimental and one control group). However, in the subgroup analysis of studies with at least 500 observations per experimental group, only eight estimates (out of 17) from five studies (out of 11) are based on sample sizes of >40,100 observations. Future work could also use multisite (multilab-style) field studies to achieve large effective sample sizes while also enabling clearer tests of heterogeneity and moderators.

Although very small effects can matter at a population scale, even nonincentivized behavioral interventions can entail significant costs (9) and scaling them requires large-scale tests to assess cost-effectiveness (7). Still, our reanalysis gives strong evidence for a nearzero average effect of nonincentivized behavioral interventions focused on citizens once publication bias is accounted for. The posterior mean is centered on zero, the credible interval is narrow, and the Bayes factor provides strong support for the null of no effect. Future research on individual and household levels should target behavioral interventions with potentially large effect sizes and should be based on adequately powered research designs to provide reliable evidence.

Our results add to the increasing review literature on behavioral interventions with and without incentives that suggests the presence of small-study effects and potentially non-negligible effect sizes even after correcting for publication bias (10). Shifting climate policy focus toward incentives on the individual level and beyond (10, 11), combined with an understanding of structural constraints (12), seems more promising than standalone behavioral interventions on citizens (13).

Materials and methods

We re-used the 144 standardized mean-difference estimates (Cohen's d) extracted by Nisa et al. from 91 studies that focus on studies without incentives (e.g. financial incentives or regulations) (4). No additional studies were added or removed.

RoBMA applies selection methods to estimate relative publication probabilities (14) and also considers small-study-effects models that relate effect sizes to their SEs, such as PET and PEESE (6, 14). RoBMA then averages across this prespecified set of candidate models to obtain a model-averaged meta-analytic effect estimate (6, 14, 15). We used RoBMA's default priors (effect: $d \sim$ Normal(0, 1); heterogeneity: $\tau \sim$ InvGamma(1, 0.15); PET $\sim$ Cauchy(0, 1) and PEESE $\sim$ Cauchy(0, 5), both truncated to $[0, \infty]$). These priors are shown to perform well across a wide range of simulation studies and empirical examples (14). Robustness of the results was assessed by rescaling the effect/heterogeneity priors (0.5×, 2×) and by a 200-iteration random one-estimate-per-study resampling analysis. Note that the reported 95% CrIs are often asymmetric with zero being the upper bound, as RoBMA combines estimates from all models based on their posterior model probabilities, introducing shrinkage toward zero if models assuming the absence of the effect predict the data best. Analyses were run in R 4.3.1 with the *RoBMA* package (16). The data and replication package contain the following sensitivity analyses: prior scaling, random one-estimate-per-primary-study resampling, and weight-function-only model.

Note

[a]Publication bias refers to selective reporting of studies based on statistical significance, whereas p-hacking refers to selective reporting of analyses within studies. From a meta-analysis perspective, these two selection processes are hard to distinguish and we refer to publication bias for brevity.

Acknowledgments

The authors are grateful to the constructive feedback of two reviewers as well as to Taisuke Imai and Tarun Khanna.

Competing Interest

The authors declare no competing interests.

Funding

F.B. acknowledges support from the Czech Science Foundation (grant #23-05227M).

Author Contributions

Adam Hardaker (Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing—original draft, Writing—review & editing), Igor Asanov (Conceptualization, Methodology, Writing—review & editing), František Bartoš (Conceptualization, Methodology, Writing—review & editing), and Stephan B. Bruns (Conceptualization, Methodology, Writing—review & editing)

Preprints

This manuscript was posted on a preprint at https://www.econstor.eu/bitstream/10419/326982/1/I4R-DP263.pdf.

Data Availability

The replication package with data and code is available at OSF: https://doi.org/10.17605/OSF.IO/FCJBE.

References

1. Dietz T, Gardner GT, Gilligan J, Stern PC, Vandenbergh MP. 2009. Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. *Proc Natl Acad Sci U S A.* 106:18452–18456.
2. Thaler RH, Sunstein CR. *Nudge: improving decisions about health, wealth, and happiness*. Yale University Press, 2008. https://doi.org/10.1016/j.soscij.2008.09.003
3. Ioannidis JPA, Stanley TD, Doucouliagos H. 2017. The power of bias in economics research. *Econ J*. 127:F236–F265.
4. Nisa CF, Bélanger JJ, Schumpe BM, Faller DG. 2019. Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change. *Nat Commun.* 10:4545.
5. van der Linden S, Goldberg MH. 2020. Alternative meta-analysis of behavioral interventions to promote action on climate change yields different conclusions. *Nat Commun.* 11: 3915.
6. Bartoš F, Maier M, Wagenmakers E-J. 2025. Robust Bayesian multilevel meta-analysis: adjusting for publication bias in the presence of dependent effect sizes [preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/9tgp2_v1
7. Asanov I, *et al.* 2023. System-, teacher-, and student-level interventions for improving participation in online learning at scale in high schools. *Proc Natl Acad Sci U S A.* 120: e2216686120. https://doi.org/10.1073/pnas.2216686120
8. Sterne JAC, *et al.* 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 343:d4002. https://doi.org/10.1136/bmj.d4002
9. Severijns R, *et al.* 2024. Involving family and friends helps sustainable diets last longer. *NPJ Clim Action.* 3:109.
10. Bergquist M, Thiel M, Goldberg MH, van der Linden S. 2023. Field interventions for climate change mitigation behaviors: a second-order meta-analysis. *Proc Natl Acad Sci U S A.* 120: e2214851120. https://doi.org/10.1073/pnas.2214851120
11. Döbbeling-Hildebrandt N, *et al.* 2024. Systematic review and meta-analysis of ex-post evaluations on the effectiveness of carbon pricing. *Nat Commun.* 15:4147.
12. Hofmann W, *et al.* 2025. Rethinking behaviour change interventions in policymaking. *Nat Hum Behav.* 9:1765–1767.
13. Imai T, Pace D, Schwarmann P, van der Weele J. 2025. Correcting consumer misperceptions about CO_2 emissions. Rationality and Competition Discussion Paper No. 529, CRC TRR 190 Rationality and Competition. https://ideas.repec.org/p/rdc/paper/529.html
14. Bartoš F, Maier M, Wagenmakers E-J, Doucouliagos H, Stanley TD. 2023. Robust Bayesian meta-analysis: model-averaging across complementary publication bias adjustment methods. *Res Synth Methods.* 14:99–116.
15. Hoeting JA, Madigan D, Raftery AE, Volinsky CT. 1999. Bayesian model averaging: a tutorial. *Stat Sci.* 14:382–417.
16. Bartoš F, Maier M, Wagenmakers E-J. 2025. RoBMA: Robust Bayesian meta-analyses [*R package*]. Version 3.5.1. Comprehensive R Archive Network (CRAN). https://CRAN.R-project.org/package=RoBMA