ARTICLE IN PRESS

International Journal of Research in Marketing xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full length article

Meta-analysis of advertising effectiveness: New insights from improved bias corrections

Joseph Korkames [a,*], T.D. Stanley [b], Stefan Stremersch [c,d]

[a] Famous Chili Inc., 1421 N 7th Street, Fort Smith, AR, USA 72901
[b] School of Business and Law, DeLMAR, Deakin University, 221 Burwood Highway, Burwood, VIC 3125, Australia
[c] Department of Marketing, Erasmus University. Contact: Erasmus University, Erasmus School of Economics, Burg. Oudlaan 50, 3062 PA Rotterdam, the Netherlands
[d] Faculty of Economics and Business Administration, Ghent University, Belgium

ARTICLE INFO

Article history:
Received 3 March 2025
Available online xxxx

Keywords:
Advertising
Sales
Meta-analysis
Publication bias
Marketing mix

ABSTRACT

The authors employ recently developed *meta*-analysis methods to better correct for misspecification, aggregation, and publication selection bias. Applying these methods in a new *meta*-analysis of advertising effectiveness, using the largest database on business-to-consumer own-brand advertising elasticities (538 elasticities), the authors show that better methods greatly reduce (more than fivefold) the *meta*-analytic overall estimate of advertising effectiveness relative to prior estimates, bringing them more in line with results from randomized controlled trials. They first obtain an average short-term elasticity of 0.0008 and a long-term elasticity of 0.03. They then explore how and when conditional advertising elasticity estimates meaningfully differ across relevant factors through advanced multiple *meta*-regression approaches that yield practically meaningful conditional estimate ranges. Over time, reported *meta*-analytic estimates of advertising-sales elasticities have become smaller, largely due to improvements in *meta*-analytic methods, while advertising effectiveness has increased—consistent with the finding that internet advertising, the most recent medium, exhibits the highest elasticity. For advertising scholars and professionals, this study provides an updated (i.e., lower) benchmark on overall advertising effectiveness, as well as more robust insights regarding when and how it varies. For analysts, it charts an improved methodological framework for marketing that should guide *meta*-analytic research on and beyond advertising effectiveness.

1. Introduction

Meta-analysis (MA) is pervasive in marketing research; it can offer robust generalizations across studies (Eisend, 2015; Hanssens, 2015, 2018). However, for MA to deliver reliable insights, it must address a fundamental challenge—*meta*-analytic bias. Without appropriate corrections, reported effect sizes may mislead both scholars and practitioners about the true impact of marketing interventions. This concern is particularly critical in advertising research, where advertising effectiveness estimates guide billion-dollar budget allocations (Kolsarici et al., 2020).

* Corresponding author.
E-mail address: joseph@famouschili.com (J. Korkames).

https://doi.org/10.1016/j.ijresmar.2025.04.001
0167-8116/

Please cite this article as: J. Korkames, T.D. Stanley and S. Stremersch, Meta-analysis of advertising effectiveness: New insights from improved bias corrections, International Journal of Research in Marketing, https://doi.org/10.1016/j.ijresmar.2025.04.001

A key debate in advertising measurement revolves around the discrepancy between ad elasticities produced by Randomized Controlled Trials (RCTs) and those from econometric models of past data (Marketing Mix Models, MMMs). Recent research argues that RCTs provide superior, unbiased estimates, while MMMs systematically overstate advertising elasticities (Gordon et al., 2019). MMM-based studies have been shown to significantly overestimate effects compared to RCT findings, leading to calls for RCTs to replace MMMs despite their higher costs and scalability limitations (Zettelmeyer, Gordon, & Bhargava, 2022).

We challenge this prevailing skepticism by demonstrating that properly conducted *meta*-analyses—when using advanced bias-correction techniques—bring MMM-based advertising elasticities much closer to those of RCTs. Specifically, we show that past *meta*-analyses of MMMs failed to correct for publication selection bias (PSB) and other distortions, leading to inflated elasticity estimates. By applying multiple *meta*-regression analysis (MMRA) with robust bias corrections, we estimate conditional effect ranges that clarify how advertising effectiveness varies across contexts. While corrected advertising effects generally range from 0.003 to 0.10, the highest elasticities (approaching 0.10) occur only under highly favorable conditions. These insights offer practical guidance for budget reallocation toward more effective channels such as digital marketing and customer relationship management (CRM), potentially increasing responsiveness and return on investment.

Beyond managerial implications, these findings resolve a major tension in advertising research. Rather than dismissing MMMs as inherently flawed, our results demonstrate that *meta*-analyses—when properly corrected—can serve as a viable and scalable alternative to expensive RCTs, while uniquely providing long-term effect estimates that experiments cannot capture.

Meta-regression analysis (MRA) presents three additional benefits. MRA (1) reveals which factors improve modeling performance; (2) shows how effects change over contextual settings; and (3) identifies the magnitude, direction, and sources of bias (Borenstein et al., 2009; Havránek et al., 2022). However, many prior *meta*-analyses in marketing fail to properly correct for *meta*-analytic bias, such as misspecification, aggregation, and PSB (Bijmolt & Pieters, 2001; Farley, Lehmann, & Sawyer, 1995; Rust, Lehmann, & Farley, 1990). While some of these issues have been partially addressed, recent advances in *meta*-analytic methods from other fields now offer improved bias-correction techniques (Nakagawa et al., 2021; Rothstein, Sutton, & Borenstein, 2006). Yet, many of these superior methods have not yet been applied in marketing and may be unknown to marketing scholars (Grewal, Puccinelli, & Monroe, 2017). Consequently, marketing researchers may be misinformed on important *meta*-analytic effects, such as advertising elasticity, because reported estimates in the literature fail to account for these biases.

To address these shortcomings, we make three key contributions:

1. Methodological Advancement: We introduce improved *meta*-analytic research methods, including an 8-step approach, to the marketing literature that better corrects for *meta*-analytic bias than prior marketing studies.
2. Revisiting advertising Elasticities: We apply these improved methods to advertising effectiveness—one of the most studied MA fields in marketing—and compare the findings to prior *meta*-analyses, demonstrating that correcting for bias brings MMM-based elasticities in line with RCT-based results.
3. Strategic and managerial Implications: By employing MMRA with multiple bias corrections, we provide new evidence on advertising effectiveness, revealing that most realistic advertising conditions yield substantially lower elasticities than traditionally reported. However, we also identify key scenarios where advertising remains highly effective, offering practical guidance on marketing strategy optimization.

Empirically, we replicate various previous MMM-based *meta*-analyses using our comprehensive database, confirming that their reported elasticities align with past literature. However, once we apply bias corrections, the short-term advertising elasticity falls to 0.0008, and the long-term elasticity to 0.03—values five times smaller than traditional estimates but consistent with RCT-based findings.

This research not only aids scholars by outlining critical variables, updating and improving heterogeneity and overall effect estimates, and investigating and revealing insights regarding relevant modeling considerations, but it also equips practitioners with tools and methodologies to tailor their advertising strategies more precisely, improving overall marketing strategy. More broadly, it offers a new perspective on the MMM vs. RCT debate, showing that properly conducted *meta*-analyses can serve as a viable and scalable alternative to expensive experiments—while uniquely offering broader insights and long-term effect estimates that RCTs cannot provide.

In Section 2, we describe common *meta*-analytic biases and methods for best practice. In Section 3, we summarize past *meta*-analyses on advertising effectiveness and detail our data collection procedures. In Section 4, we report our findings. In the discussion (Section 5), we detail the sources and impact of bias in the advertising literature, implications for both scholars and practice, the limitations of our research, and future research directions.

2. Meta-analytic biases and modeling methods

In this section, we review state-of-the-art MA methods before we turn to our empirical application. We categorize common biases in *meta*-analyses into three types. We then detail and compare conventional MA weighted averages. Next, we describe how to correct MA weighted averages using MRA to improve overall average effect estimates. Then, we describe MMRA in conjunction with misspecification, aggregation, and publication selection corrections to estimate more comprehensively corrected conditional effect estimate ranges. Lastly, we summarize the 8-step *meta*-analytic modeling approach we advocate and employ in the empirical context examined-advertising effectiveness.

2.1. Common biases in meta-analysis

First, misspecification biases are induced by improper modeling, overfitting, and/or omitted variables (Greene, 1990; Gujarati, 1995). MA overall averages are simple weighted averages. However, if there is a systematic pattern among these reported effects that depends on, for example, which variables are included in the primary studies, or whether endogeneity is accounted for, then these averages will be biased due to these omitted variables.

Secondly, aggregation biases are induced by aggregating incomparable variable measures (e.g., apples versus oranges) at the MA stage (Sharpe, 1997). For example, a study may measure advertising effectiveness either by the increase in sales revenue or the change in the market share of the product. If both are lumped together, the estimate is an average of these separate effects whose magnitude may potentially differ, depending on the relative mix.

Thirdly, PSB includes traditional reporting bias—the preference of some researchers and/or reviewers to report statistically significant effect estimates—as well as other selection biases from *p*-hacking and questionable research practices (Eisend & Tarrahi, 2014; Simmons, Nelson, & Simonsohn, 2011; Lin et al., 2020). PSB often leads to exaggerated findings being reported in scientific literature (Sterling, 1959; Rothstein et al., 2006; Moreno et al., 2009).

2.2. Random and fixed effect estimators

The most common *meta*-analysis estimators are random-effects (RE) and fixed-effect (FE) weighted estimators (Borenstein et al., 2010). Random effects estimators assume that the reported effects, y_i, are randomly and normally distributed around the mean, μ, with i representing an observation instance and k the number of estimates collected. Random effects is the most frequently used MA estimator across disciplines as it allows effects to vary (i.e., heterogeneity, θ) beyond random sampling errors, ε, alone (Borenstein, Hedges, & Rothstein, 2007). Namely,

$$y_i = \mu + \theta_i + \varepsilon_i \quad i = 1, 2, \ldots, \mathrm{k} \tag{1}$$

Random effects MA does not require that all 'true' estimates be the same, as assumed by fixed-effect *meta*-analysis. Rather, 'true' effects, $\mu + \theta_i$, are assumed to vary randomly around the population mean elasticity, m.

All *meta*-analysis estimators are simple weighted averages with optimal inverse-variance weights that down weight smaller studies (i.e., those with larger standard errors), which also tend to have larger biases, thereby decreasing publication bias when present (Borenstein et al., 2007; Danaher & Brodie, 2000). These standard errors must be gathered or calculated based on the estimating model in the relevant primary studies. For the random effects estimator, these inverse-variance weights are $\frac{1}{SE_i^2+\hat{\tau}^2}$, where SE_i^2 is the square of each estimate's standard error, SE, and $\hat{\tau}^2$ is the estimated heterogeneity variance (i.e., the variance of θ_i).

The fixed effects estimator uses weights of $\frac{1}{SE_i^2}$, as it assumes that there is no heterogeneity among reported effects but only random sampling error. Both are examples of weighted least squares (WLS) widely used in regression to adjust for observed differences in the variance of the dependent variable, y_i, (i.e., heteroskedasticity).

2.3. Unrestricted weighted least squares meta-analysis

Like random effects, unrestricted weighted least squares (UWLS) adjusts for heterogeneity in true effects; its weights are $\frac{1}{\gamma SE_i^2}$ for $\gamma > 0$ (Stanley & Doucouliagos, 2015). Unlike random effects, UWLS allows the heterogeneity variance to be correlated with SE_i^2, as has been found to be the case in medical research and psychology (IntHout et al., 2015). Empirically, UWLS has also been shown to better fit data over a multitude of research fields compared to both random effects and multilevel methods (Mawdsley et al., 2017; Yang et al., 2023). Furthermore, UWLS has been shown to be less upward biased than random effects under a wide range of simulation settings when publication selection bias is present (Bom & Rachinger, 2019; Henmi & Copas, 2010; Stanley & Doucouliagos, 2015).

UWLS always has the same point estimate as fixed effects but with larger standard error and wider confidence intervals (CIs) than fixed effects when there is heterogeneity. UWLS is calculated by a simple regression of the standardized effect size (y_i/SE_i) on its precision ($1/SE_i$) with no intercept.

$$t_i = y_i/SE_i = \alpha(1/SE_i) + u_i \quad i = 1, 2, \ldots, \mathrm{k} \tag{2}$$

The UWLS estimate is given by the estimated slope, $\hat{\alpha}$. We suggest authors rely upon UWLS but also report random effects, as we do here.

2.4. Funnel asymmetry test – precision effect test – precision effect estimate with standard error (FAT-PET-PEESE)

FAT-PET-PEESE is a *meta*-regression approach for modeling publication selection bias estimated using UWLS that additionally includes the standard error as a regressor term and estimates effects conditional on $SE = 0$ (Stanley & Doucouliagos, 2012). When publication selection bias is not present, the effect (e.g., an elasticity estimate), y_i, and its standard error, SE_i, will be uncorrelated. In contrast, a positive correlation between the magnitude of y_i and the size of SE_i by a standard *t*-test on β_1, suggests that publication selection bias or a small-study effect is present (Egger et al., 1997; Stanley & Doucouliagos, 2012). FAT-PET-MRA (Equation (3)) is estimated using WLS with $1/SE_i^2$ as the weights.

$$y_i = \beta_0 + \beta_1 SE_i + \varepsilon_i \tag{3}$$

FAT tests $H_0 : \beta_1 = 0$. The rejection of H_0 indicates funnel asymmetry, which can be interpreted as publication selection bias when heterogeneity is suitably addressed via the inclusion of relevant moderators in MMRA models and when its results are corroborated by alternative PSB methods, such as those we mention below. The $\beta_1 SE_i$ term represents PSB, and a statistically significant positive β_1 coefficient on the standard error term (FAT) in Equation (3) can be interpreted as selection for statistically significant positive effects when the above conditions are met. With these WLS weights but without the $\beta_1 SE_i$ term, Equation (3) reduces to UWLS.

PET tests $H_0 : \beta_0 = 0$, the rejection of which provides evidence for the presence of a genuine effect beyond publication selection bias. When there is evidence of an effect beyond PSB, PEESE is a better estimate of its magnitude. The MRA model for PEESE is the same as Equation (3), except that SE_i^2 replaces SE_i as the independent variable (Stanley & Doucouliagos, 2012). The constant term, β_0, in Equation (3) represents the PSB-corrected effect.

While it assumes a linear relationship between effect sizes and their standard errors, research indicates that this assumption is generally robust. For instance, Alinaghi and Reed (2018) found that PET-PEESE performs well in various scenarios, maintaining nominal false-positive rates even under typical conditions encountered in *meta*-analyses. Additionally, Almalik (2024) demonstrated that PET-PEESE remains effective in the presence of heteroscedasticity, suggesting its robustness across different data conditions.

We use and recommend FAT-PET-PEESE as a principal methodology as it has been shown to work well in practice and can be extended to multilevel multiple *meta*-regression (Nakagawa et al., 2021). However, we recommend cross-validating FAT-PET-PEESE results with at least two alternative methods, such as the 3-parameter selection model (3PSM) and the Test of Excess Statistical Significance-Proportion of Statistical Significance Test (TESS-PSST). We report these in Web Appendices H and J for our empirical application (Bartoš et al., 2022; Carter et al., 2019).

2.5. Bias-corrected multiple meta-regression analysis (MMRA) conditional estimation

The above methods are useful for improving overall averages and testing for PSB. Here we detail how FAT-PET PEESE can be combined with MMRA to further calibrate conditional estimates to more comprehensively and simultaneously account for misspecification, aggregation, and publication selection.

With MMRA, analysts can explain the wide variation in reported effects, estimate how the mean effect varies across relevant conditions, and allow for the estimation of effects under the best research practices, i.e., provide more robust conditional estimates than MA / MRA overall effects.

The MMRA model is:

$$y_i = \beta_0 + \sum_j \beta_j X_{ij} + \varepsilon_i \quad \mathrm{i} = 1, 2, \ldots, \mathrm{k} \quad \& \quad j = 1, 2, \ldots, \mathrm{J} \tag{4}$$

where X_{ij} are independent variables capturing conditions under which an effect is estimated, and J is the number of variables included. We recommend that *meta*-analysts code variables based on those found relevant by past primary studies and *meta*-analyses, and estimate the above model using UWLS with $1/SE_i^2$ as the weights. For the sake of robustness, we also report the random-effect MMRA results in the Web Appendix, see Table WH5. Random effects *meta*-regression adds another random heterogeneity term, θ_i, into Equation (4) and estimates its variance separately. UWLS-MMRA, in contrast, estimates the composite sampling error and heterogeneity variance together. We suggest analysts incorporate study-level cluster-robust SEs in all MMRA models to account for potential dependencies across elasticities within studies that can bias MMRA SE.

To further accommodate dependence within studies, we suggest multilevel (or panel) models:

$$y_{is} = \beta_0 + \sum_j \beta_j X_{isj} + \vartheta_s + \varepsilon_i \quad \mathrm{i} = 1, 2, \ldots, \mathrm{kj}; \quad \mathrm{j} = 1, 2, \ldots, \mathrm{J}; \quad \mathrm{s} = 1, 2, \ldots, \mathrm{S} \tag{5}$$

with UWLS weights of $1/SE_i^2$ and where ϑ_s are the study effects, which can be random or fixed, with S representing the number of studies included in the analysis. We recommend analysts estimate this multilevel model using fixed effects unbalanced panel methods. Fixed effects panel models account for any unobserved study effects (including any unobserved differences in study 'quality') not accounted for using clustering or the X_{isj}-variables. As a result, it elevates 'observational' statistical analysis with all of its threats to validity to a stronger quasi-experimental design (Rockers et al., 2015).

However, multicollinearity can be an increased concern when including and testing many moderators. Thus, we suggest a general-to-specific variable selection procedure to first simplify Equation (4), where possible, letting the research record objectively select the more important variables (Campos, Ericsson, & Hendry, 2005).

To obtain bias-corrected MMRA estimates, analysts need to estimate conditional effects by specifying the values of X_{isj} that represent the best research settings (i.e., those without misspecification and PSB) or the more context-relevant settings (aggregation bias corrections). This allows analysts to provide an effect estimate, or range of potential effects, based on relevant linear combinations of X_{isj}, where aggregation-related variables may be varied to establish a range, when applicable—see Section 4, Table 3.

2.6. Meta-analytic inference: an 8-step approach

We recommend an 8-step approach to test for *meta*-analytic bias and provide PSB-corrected overall effects and comprehensively corrected conditional effect estimates using MMRA, which we apply below in our MA on adverting effectiveness. This method starts with first building an understanding and depictions of the data, then applying gradually more sophisticated methods to better understand the data and effect under varied conditions and accounting for different factors before using the most robust, literature-supported methods to estimate a range of relevant conditional effects across relevant settings and effect heterogeneity.

In Step 1, we provide descriptive distributional statistics and graphics including: the unadjusted mean, median, funnel plots, and primary study box-plots as a baseline. In Step 2, we employ RE to estimate the overall average effect, examine the magnitude of $\hat{\tau}^2$, and compare with the descriptive statistics.

In Step 3, we calculate UWLS with cluster-robust standard errors (Equation (3) and forgo FE, as it is equal to UWLS but with narrower confidence intervals. In Step 4, we employ FAT-PET-PEESE by adding the standard errors as an independent variable with UWLS, $1/SE_i^2$ weights. At this point, we cross-validate the effect estimates and presence of PSB found with subgroup analyses and alternative methods—at least 3PSM and TESS-PSST. We detail all alternative methods and subgroups used in the robustness checks in Section 4 and Web Appendices I – K.

In Step 5, we estimate a comprehensive MMRA model by adding moderators to the FAT-PET MRA model, following Equation (4), and test for systematic heterogeneity. In Step 6, we apply general-to-specific to simplify the UWLS FAT-PET MMRA model to those factors that provide statistically meaningful differences, arriving at a specific model, following Equation (4). Here, we can also estimate the RE model for cross-validation.

In Step 7, we include study-level effects, following Equation (5), and estimate a fixed effects panel-specific model from Step 6. For each model in Steps 5 – 7, we provide a range of effects that can be expected across typical conditions by varying variables associated with aggregation bias and setting the specific values of those variables associated with misspecification and PSB to represent best research practices.

In the final step (Step 8), we examine the overall results across models and conditions, and use the models from Steps 5 – 7 to establish a generalizable range and/or to estimate relevant effects for the contexts and conditions of interest.

2.7. Recent and best practices: findings of the meta-analysis methods literature

Above we detailed how *meta*-analyses should test and correct for common *meta*-analytic biases including misspecification, aggregation and publication-selection bias. However, what are studies currently doing, and how does that compare to what they should be doing? Below we detail recent (or current) versus best research practices for conducting *meta*-analyses with reference-backed statements that, taken together, show the superiority of our approach compared to current practice.

Misspecification biases: recent practices. MA weighted averages (e.g., RE) cannot account for misspecification. While 'meta-analyses' in econometric-based marketing studies have used multiple regression to examine heterogeneity, prior advertising effectiveness studies have not estimated elasticities from multiple regression models in a conditional way that controls for better practice or avoids identified biases (Henningsen et al., 2011; Sethuraman et al., 2011).

Misspecification biases: best practices. Among other things, MMRA needs to include independent variables that track whether potentially relevant independent variables were included in the primary study's regression model (Stanley & Doucouliagos, 2012, p. 3). By coding dichotomous (0/1) variables for the presence/absence of relevant variables that would cause bias if omitted, misspecification biases can be corrected or reduced. Studies further need to control for reverse causation (endogeneity) and for unobserved differences in markets and products, by utilizing multilevel (or panel) data (Stanley & Doucouliagos, 2012). To reflect these best practices, the *meta*-analyst can calculate effects by including independent variables that indicate whether the primary literature itself includes potentially relevant independent variables, employs causal methods, or employs better or more informative data. This allows us to estimate conditional effects under widely accepted best research settings.

Aggregation biases: recent practices. Aggregation bias can stem from natural heterogeneity in effect across contexts, or from improperly aggregating incomparable phenomena. The problem of comparing *apples and oranges* has long been discussed and addressed by the MA literature (Sharpe, 1997). Eysenck (1984) famously noted that, "Adding apples and oranges may be a pastime of children learning to count, but unless we are willing to disregard the differences between these two kinds of fruit, the [*meta*-analytic] result will be meaningless." Practitioners must realize that only strictly *comparable* estimates can be meaningfully combined. Unfortunately, theoretically different effects are sometimes inappropriately aggregated, however, there has been attempts to alleviate these issues still. While recent predecessor studies do not significantly disaggregate estimates, they do disaggregate by focusing on B2C brand-level estimates, not including B2B or category elasticities, which we follow in our main analyses.

Aggregation biases: best practices. To minimize possible aggregation biases, effects need to represent the same phenomenon and be measured in the exact same units (e.g., correlation coefficients or the own-sales elasticity of advertising). To allow for potentially more nuanced differences within comparable estimates, independent variables and subgroup analyses can be employed. For example, own-sales advertising elasticities and cross-sales elasticities (the effect of advertising on a competitor's sales) are distinct phenomena and should not be combined in the same *meta*-analysis. MMRA offers objective statistical approaches that allow the research record to determine if these measures induce meaningfully different effect sizes.

Publication and selection biases: recent practices. Marketing and behavioral research has often historically followed versions of Schmidt and Hunter (2015) to correct for study artifacts. However, these methods are more relevant for *meta*-analyses of correlations as opposed to econometric estimates, as is our focus.

There are several MA techniques available to test and/or correct for publication and selection biases. Several are widely known to perform poorly (Nakagawa et al., 2021). Unfortunately, marketing research often still employs discredited methods: Fail-Safe N, Trim and Fill, and the P-Curve (Gelman, 2018; Nakagawa et al., 2021; Rothstein et al., 2006; Ruan, Hsee, & Lu, 2018; van Aert, Wicherts, & van Assen, 2016).

Variations of Rosenthal (1979)'s Fail-Safe N are still employed in marketing (Roschk & Hosseinpour, 2019; Rosengren et al., 2020). Fail-Safe N calculates the number of null effect studies that would need to be added to reduce a statistically significant overall MA result to non-significance (Rosenberg 2005, p. 464). Becker in Rothstein et al. (2006) argues Fail-Safe N should be abandoned as it is "prone to misinterpretation and misuse and no statistical criterion is available for interpretation" (p. 111).

Trim and Fill is still used to test for, and correct for, publication selection bias in marketing by adjusting the *meta*-analytic estimate for the potential impact of missing studies. Trim and Fill works by examining a funnel plot for asymmetry, with the assumption that this asymmetry is due to suppressed nonsignificant results. This symmetry assumption need not be valid when there is heterogeneity, such as for advertising elasticities. Trim and Fill performs poorly in both applications and simulations (Bartoš et al., 2022; Carter et al., 2019; Nakagawa et al., 2021).

P-curve provides a test for whether an effect would still be detected if p-hacking were not present. P-hacking is the use of any number of questionable research practices (QRPs): selectively reporting only some of the outcome variables collected; stopping data collection or estimation when a statistically significant result is obtained; or rounding p-values down. P-hacking invalidates the Fail-Safe N (Simonsohn, Nelson, & Simmons, 2014 a, b). The P-curve looks only at the distribution of statistically significant *p*-values. When there is no effect and no publication bias, the distribution of *p*-values is uniform. When there is a genuine non-null effect, this distribution becomes right-skewed, which can serve as a test of whether a true effect is present and can provide a publication selection-corrected estimate (Carter et al., 2019). However, P-curve can overestimate effects; it is no longer recommended for use for multiple reasons highlighted in Carter et al. (2019); McShane, Böckenholt, & Hansen (2016); and van Aert et al. (2016).

Marketing has developed methods for measuring the size of publication bias, such as in Rust et al. (1990); however, such methods did not carry forward either in marketing or otherwise.

Publication and selection biases: best practices. Best practices for testing and/or correcting for publication selection bias tend to fall into two categories: selection models such as 3PSM, or regression-based methods such as FAT-PET-PEESE. These models broadly reduce the bias of MA weighted averages and mean squared error (Bartoš et al., 2022; Carter et al., 2019; Iyengar & Greenhouse, 1988; Nakagawa et al., 2021; Rothstein et al., 2005).

More recently, TESS-PSST has shown to be a more powerful test for the presence of publication selection bias. TESS-PSST combines two tests: TESS (Test of Excess Statistical Significance) and PSST (Proportion of Statistical Significance Test) (Stanley et al., 2021). PSST investigates the difference between the observed and expected proportions of statistically significant findings through a simple test of a proportion. The expected proportion of statistical significance is calculated as if there is no selection for statistical significance from the MA estimate of the mean effect and each estimate's standard error. When the observed proportion is significantly larger than the expected proportion, this suggests publication selection bias.

TESS is a simple test of the proportion of excess statistical significance against the null hypothesis that it is 5 %. See Stanley et al. (2021) for details, codes, and calculations. If either TESS or PSST is statistically significant, this is a signal of publication selection bias. Together, TESS-PSST dominates 3PSM and FAT under various settings (Stanley et al., 2021). We recommend including TESS-PSST in tests to identify publication selection bias. Unlike FAT, a significant result cannot be dismissed as a 'small-study' effect; TESS-PSST makes no assumption about the correlation of effects and their standard error or treat small studies differently.

Simulation studies have suggested FAT-PET-PEESE, 3PSM, and TESS-PSST are among the preferred methods to accommodate or test for publication selection bias. However, 3PSM and other selection methods, at times, do not converge. Regardless, no publication selection method comes without fault, or is optimal across all settings (Bartoš et al., 2022; Carter et al., 2019).

Some suggest using multiple publication bias approaches and looking for robustness among them (such as Carter et al. (2019)), while Bartoš et al. (2022) suggest a Bayesian model average of several methods is preferrable. We use and recommend the use of FAT-PET-PEESE as a principal methodology, as it has been shown to work well in practice and can be extended to multilevel multiple *meta*-regression (Nakagawa et al., 2021). However, we also employ 3PSM, RE, TESS-PSST, UWLS, and WAAP (Weighed Average of the Adequately Powered Studies) to corroborate our findings and to look for robust agreement across methods. Method summaries and method comparisons can be found in: Bartoš et al. (2022); Carter et al. (2018); Nakagawa et al. (2021); Rothstein et al. (2005); Stanley and Doucouliagos (2015); and van Aert et al. (2016). For broader best practices and guidance that complement those presented here, see Korkames (2024a).

The presence of biases combined with insufficient use of state-of-the-art methods, some of which have been recently developed after the publication of past *meta*-analyses in marketing, has plagued the prior studies, potentially leading to inflated estimates as we later show is the case in advertising (see Tables 2 – 3 in Section 8). However, the above best practices can largely correct these biases. We summarize recent versus best practices via Table WA2 in Web Appendix A.

3. Prior research on advertising effectiveness and meta-analytical sample design

In this section, we first review past *meta*-analyses in advertising, including the methods they use, their key findings, and their limitations. We then describe our data and coding.

3.1. Prior meta-analytic research on advertising effectiveness

To ground our focus on advertising, we provide the current state of this research area, and detail the importance of undertaking this particular case (advertising effectiveness). Firms spend more on advertising than almost any other marketing activity (Horsky, Horsky, and Zeithammer 2016). However, the wide variation in findings across studies with diverse measures and methods has made generalizable conclusions elusive. As such, MA is often employed in this research area to consolidate and determine robust findings.

Advertising effectiveness has been the subject of some of the earliest *meta*-analyses in marketing (Clarke, 1976; Lodish et al., 1995) and continues to attract *meta*-analyses (Eisend, 2009; Edeling & Fischer, 2016; Köhler et al., 2017; Rosengren et al., 2020; Schöndeling et al., 2023). Marketing studies commonly evaluate advertising effectiveness as elasticities by measuring the percent change in sales, market share, and/or consumer choice that is associated with a 1 % change in some measure of advertising spending (Korkames and Stremersch, 2024).

In marketing, three *meta*-analyses have focused on measuring overall empirical advertising elasticities (Assmus et al., 1984; Henningsen et al., 2011; Sethuraman et al., 2011). To establish the research baseline, we review some of their key findings below.

The first MA estimating advertising elasticities finds smaller elasticities for: time-series data, the US as compared to Europe, non-food products, and in the short-term (Assmus et al., 1984). Although Assmus et al. (1984) does mention publication bias and states that the inclusion of unpublished estimates would be "very desirable in the future", the authors do not incorporate them. Similarly, they correctly mention (mis)specification bias as "caused by omission of variables correlated with those included in the [primary study] equation" (p. 67) but do not adjust for these biases. From the beginning, marketing researchers have acknowledged the importance of misspecification and publication selection bias in *meta*-analysis.

Sethuraman et al. (2011) provide simple averages and examines heterogeneity via multilevel regression. They find that advertising effectiveness has decreased over time and that advertising elasticities are higher for durable goods, early in the life cycle, in yearly compared to quarterly data, and when advertising is measured in gross rating points (GRP).

The smallest *meta*-analytic own-brand elasticity estimates reported in the marketing literature are 0.09 (short-term) and 0.19 (long-term) (Henningsen, et al., 2011). They find advertising is higher for hedonic and experience goods, for new goods, and when advertising is measured in GRP. Like Sethuraman et al. (2011), and following Bijmolt and Pieters (2001), Henningsen and colleagues report unadjusted, simple averages for overall effect estimates and employ multilevel, multiple regression methods to describe heterogeneity. For more details regarding the findings from these studies, see Web Appendix A and Table WA1.

Methodologically, past *meta*-analyses of advertising elasticities (Assmus et al., 1984; Henningsen et al., 2011; Sethuraman et al., 2011) do not employ MA weighted averages for overall effect estimation or these more sophisticated *meta*-regression methods described above. They do not test, or correct for, misspecification, aggregation, or publication selection bias in their estimates provided. They do use multilevel regression to describe heterogeneity, but without inverse-variance weights as routinely employed in random effects, fixed effects, and UWLS MA and *meta*-regression, without study-level cluster-robust standard errors leading to inaccurate heterogeneity insights, and without incorporating the standard error as an independent variable in the regression leading to oversized estimates. Furthermore, they do not provide conditional estimates. Additionally, they incorrectly assume that the reported effect is independent of its standard error, making their *t*-tests examining the effect of advertising on sales invalid.

Previous advertising elasticity *meta*-estimates, though decreasing over time, have estimated overall average effects between 0.09 and 0.47. While *meta*-analytic methods used in advertising have clearly been improving, advertising *meta*-analyses have yet to employ conventional MA weighted averages, or conditional estimates, or test and accommodate for publication selection bias, widely acknowledged as problematic for *meta*-analyses (Borenstein et al., 2009). However, accurate estimates of advertising effectiveness and heterogeneity are imperative for firms' marketing allocations and sales growth forecasting (van Heerde et al., 2013). In the next sub-section, we summarize the primary study literature in this area.

3.2. Primary study literature on advertising effectiveness

The primary studies that estimate advertising effectiveness are extremely varied. Researchers have used various methodologies to understand how advertising influences consumer behavior, brand recognition, and sales such as those presented in Lavidge and Steiner (1961); Batra, Myers, and Aaker (1996); Puto and Wells (1984); Lodish et al. (1995); Lewis and Reiley (2014); Naik and Raman (2003); and MacKenzie and Lutz (1989). These are presented in more detail in Web Appendix A.

3.3. Data collection and coding protocols

In the following sub-section, we describe our data collection and coding protocol. For a compact summary of the prior *meta*-analyses compared to ours, distinct but related articles, and the theoretical motivation for the variables we use, see Table WA1 and Web Appendix A.

Our data consists of advertising sales effectiveness elasticities (dependent variable) and dozens of factors that might influence the magnitude of this elasticity (independent variables).

We initially worked with a private analytic team to collect and code the relevant advertising elasticity literature and standard errors (SE)—935 advertising elasticity estimates and 25 independent variables from 74 studies. We extracted 538 B2C own-brand estimates from 40 empirical studies that contained nonzero standard errors required for our main sample to be analyzed with the methods we advocate. We provided *meta*-analytic weighted averages and publication selection corrections in Table 1, compared to previous *meta*-analyses in Table WA1, and estimated alternative subgroups and models compared to Table 1 in Tables WH1-WH4. We then presented our main results in Table 2, with alternatives to Table 2 presented in Table WH5 and Web Appendices I–J. After completion of this initial work, we later collected additional estimates for a final set of robustness checks making the total number of points analyzed up to 1096 estimates from 93 papers in Tables 4 – 5.

3.3.1. Study inclusion

We include all studies that provide empirical, brand-level, business-to-consumer, own-sales advertising elasticities with retrievable, nonzero standard errors. We exclude (1) effects based on noneconometric methods such as those in Lodish et al. (1995); (2) cross-elasticities; (3) effects not convertible to elasticities; (4) outcomes not related to sales, market share, or choice; and (5) estimates without standard errors or related statistics. As in prior *meta*-analyses, we include various outcome (e.g., monetary sales, choice, and market share) and input measures (e.g., monetary advertising, relative advertising, and GRP), but we use independent variables, subgroup analyses, and multilevel MMRA analyses to account for distinct outcome and input measures (for further details of the inclusion criteria and coding see Web Appendices B – F).

3.3.2. Search methodology and data acquisition

Independent coders searched Google Scholar, Science Direct, and Web of Science for references containing "advertisement" or "advertising" and "elasticity" in the title, abstract, or keywords. They also searched NBER working papers for "sales

Table 1
Meta-analysis and meta-regression: overall average effects.

Short-term[c]	(Model 1) Prior Research: Simple Average[a]	(Model 2) RE Meta-Analysis	(Model 3) UWLS Meta-Analysis	(Model 4) FAT-PET MRA[b]
Elasticity	0.09	0.02* (0.001)	0.004* (0.002)	0.0008 (0.001)
SE	–	–	–	1.90* (0.3)
n	496	496	496	496
Long-term [c]				
Elasticity	0.14	0.08* (0.005)	0.05* (0.02)	0.03 (0.03)
SE	–	–	–	1.86 (0.92)
n	42	42	42	42

[a]Replicates Sethuraman et al. (2011) and Henningsen et al. (2011)'s simple average using our sample. We find the same short-run effect as Henningsen et al. (2011) and a qualitatively similar long-term effect in our sample using a simple average to replicate the method of prior studies.
[b]We consider FAT-PET MRA in Model 4 best practices for overall MA estimates but prefer the MMRA model conditional estimates in Table 2.
[c]These estimates are simple *meta*-analytic weighted averages. See Table 2, our main results after accounting for dozens of effects and biases in estimating advertising elasticities.
Notes. Numbers in parentheses are cluster-robust standard errors. * indicates $p < 0.05$.

Table 2
Multiple meta-regression models and conditional estimates.

Independent Variable	**(Model 1) General UWLS MRA**[a]	**(Model 2) Specific UWLS MRA**	**(Model 3) Specific FE PANEL MRA**
Constant (Model Intercept)	0.001 (0.003)	–0.002 (0.003)	0.009* (0.005)
SE	1.42* (0.22)	1.64* (0.28)	1.75* (0.05)
Published	–0.003 (0.002)	–0.002* (0.0007)	–0.001* (0.0005)
Annual	0.01* (0.002)	0.01* (0.002)	0.01* (0.0005)
Full Marketing Mix	–0.008* (0.004)	–0.009* (0.004)	–0.005* (0.001)
Abs: GRP	0.02 (0.02)	0.027* (0.009)	0.008 (0.005)
Competition	–0.008 (0.005)	–0.004* (0.002)	–0.003* (0.001)
Growth	0.03* (0.002)	0.03* (0.002)	0.02* (0.0006)
Omit: Lag Price	0.008* (0.003)	0.004* (0.002)	–0.008 (0.004)
Entertainment Media	0.05* (0.005)	0.05* (0.004)	0.002 (0.002)
Panel	0.0007 (0.0004)	0.002* (0.0007)	0.001* (0.0002)
US	0.009* (0.002)	0.006* (0.0007)	0.004* (0.0007)
TV	0.008* (0.003)	0.09* (0.007)	0.004* (0.001)
Internet	0.07* (0.009)	0.01* (0.002)	0.05* (0.002)
Display	0.008 (0.006)	–	–
Dynamic	0.001 (0.002)	–	–
Dep: Market Share	–0.0004 (0.004)	–	–
Abs: Monetary Spending	–0.003 (0.002)	–	–
Omit: Lag Advertising	0.0007 (0.003)	–	–
Durables	0.04 (0.05)	–	–
Long-term	0.03 (0.02)	–	–
Endogeneity	0.002 (0.04)	–	–
Time Trend	–0.003 (0.004)	–	–
Dep: Choice	0.003 (0.002)	–	–
Seasonality	–0.002 (0.001)	–	–
Recession	–0.02 (0.06)	–	–
Conditional Estimate Range [b]	[-0.04, 0.13*]	[-0.01*, 0.15*]	0[.003, 0.10*]

[a]We look across models. See Web Appendix F and this work for more details on effect estimation.
[b]We estimate numerous conditional estimates; context-specific effects vary more than our conditional range. See Web Appendix F and our included code for details of how we estimate both.
Notes. Numbers in parentheses are cluster-robust standard errors. * indicates $p < 0.05$. n = 538 estimates. S = 40 studies.

elasticity" or "advertising elasticity". We specified that studies were to be considered relevant if they measured a comparable advertising effect and sufficient information was reported to retrieve or calculate standard errors, such as a *t*-test. In Web Appendix F, we provide a PRISMA diagram (see Figure WE1), our search method (see Table WE1), and the studies we included with the distributions of the elasticities they report (Web Appendix G). The first author conducted forward and backward citation searches of Assmus et al. (1984) and Sethuraman et al. (2011) to identify additional relevant MA and emailed all identified authors requesting their data. The only author team that responded positively to our request and shared their data was Henningsen et al. (2011). We re-coded, checked, and augmented the received data in terms of variables and studies included. In total, we compiled 935 estimates from 74 studies (538 estimates from 40 studies had retrievable, nonzero SE). We did not contact the authors of Clarke (1976) or Köhler et al. (2017), given different focus.

3.3.3. Coding protocols and data preparation

The team coded advertising elasticities (including their standard errors) and 25 independent variables from all 40 relevant studies (see Web Appendices A and F and Table 3 for variable definitions and theoretical motivations and groupings). We detail the relationship of each variable to potential misspecification, aggregation, and publication selection bias in Table 3. Our team's coding was in more than 96 % agreement with the data received from Henningsen et al. (2011). Disagreements between coders were resolved by consulting the primary studies and emailing authors.

In total, we obtained 935 advertising elasticity estimates from 74 studies, of which 538 B2C own-brand estimates from 40 empirical studies had retrievable nonzero standard errors for our main sample (see the data on WEBSITE). We provide *meta*-analytic weighted averages and simple publication selection corrections in Table 1, compare to previous *meta*-analyses in Table WA1, and estimate alternative subgroups and models compared to Table 1 in Tables WH1-WH4. We then present our main results in Table 2, with alternatives to Table 2 presented in Table WH5 and Web Appendices I –J. After completion of this initial work, we collected additional estimates for further robustness checks, bringing the total number of points analyzed to 1,096 estimates from 93 papers. We present those robustness checks in Tables 4 – 5.

Table 3
Variable definitions, bias categories, and conditional values.

Variable Name	Variable Definition	Bias Types Misspecification	Aggregation	Publication Selection	Characteristic Type	Ideal values to calculate *conditional estimates
Elasticity	DV coefficient after converted to an elasticity.	–	–	–	NA	NA
SE	The standard error for the elasticity coefficient.	X		X	PSB	0
Published	1 if published paper, 0 otherwise.			X	PSB	0
Annual	1 if the study uses annual data, 0 otherwise.		X		Data	Depends
Full Marketing Mix	1 if the full marketing mix is considered, 0 otherwise.	X			Modeling	1
Abs: GRP	1 if advertising is measured in gross rating points, 0 otherwise.		X		Data	Depends
Competition	1 if model accounts for attributes of competition, 0 otherwise.	X			Modeling	1
Growth	1 if estimate is from a growth product (life cycle), 0 otherwise.		X		Firm	Depends
Omit: Lag Price	1 if the model omitted the lagged price, 0 otherwise.	X			Modeling	0
Entertainment Media	1 if product is entertainment media, 0 otherwise		X		Firm	Depends
Panel	1 if a panel data structure is used.	X			Data/Modeling	1
US	1 if the country is the US, 0 otherwise.		X		Firm	Depends
TV	1 if the estimate is from a tv or movie ad, 0 otherwise.		X		Firm	Depends
Internet	1 if the estimate is from internet ads, 0 otherwise.		X		Firm	Depends
Display	1 if the estimate is from a display ad, 0 otherwise.		X		Firm	Depends
Dynamic	1 if lagged dependent variable is included (any lags), 0 otherwise.	X			Modeling	1
Dep: Market Share	1 if the estimate is a market share estimate, 0 otherwise.		X		Data	Depends
Abs: Monetary Spending	1 if advertising is measured as absolute monetary value, 0 otherwise.		X		Data	Depends
Omit: Lag Advertising	1 if the model omitted the lagged ad spending, 0 otherwise.	X			Modeling	0
Durables	1 if the estimate relates to a durable good, 0 otherwise.		X		Firm	Depends
Long-term	1 if estimate is for long-term as detailed by study, 0 otherwise.		X		Modeling	Depends
Endogeneity	1 if study controls for endogeneity (IV or GMM), 0 otherwise.	X			Modeling	1
Time Trend	1 if a time trend is included, 0 otherwise.	X			Modeling	1
Dep: Choice	1 if DV measured as a choice outcome, 0 otherwise.		X		Data	Depends
Seasonality	1 if seasonality is controlled for, 0 otherwise.	X			Modeling	1
Recession	1 if seasonality is controlled for in the model, 0 otherwise.	X			Modeling	1

*These values are used to calculate conditional estimates for the mean advertising elasticity in the results of the MMRA estimation in Table 2

4. Empirical analyses and results

In this section, we conduct the 8-step method that we suggest in Section 2.6.

4.1. Descriptive statistics

We display the distribution of short and long-term advertising elasticity data via the funnel plots in Figs. 1 – 2 and the histograms in Fig. 3 – 4, respectively. For detail regarding how short-term and long-term are defined here, see Web Appendix F. The solid lines represent the median, and the dashed lines denote the mean. The distributions are highly right-skewed,

Table 4
MA: aggregate.

Short-term[a]	**RE Meta-Analysis (1)**	**UWLS Meta-Analysis (2)**	**RE MA with SE (3)**	**FAT-PET MRA (4)**	**FE PANEL MRA with SE (5)**
Elasticity	0.03* (0.0008)	0.005* (0.002)	0.007* (0.003)	0.001 (0.001)	0.0005 (0.0009)
SE	–	–	1.35* (0.08)	1.60* (0.19)	1.65* (0.07)
n	1004	1004	1004	1004	1004
Long-term[a]					
Elasticity	0.04* (0.003)	0.01* (0.006)	0.0004 (0.009)	0.0008 (0.004)	0.003 (0.014)
SE	–	–	1.83* (0.25)	1.86* (0.49)	1.75 (0.88)
n	92	92	92	92	92

Notes. Numbers in parentheses are cluster-robust SE except RE. * indicates $p < 0.05$.
[a]These estimates are simple *meta*-analytic weighted averages. See the MMRA models for main results.

Table 5
Multiple meta-regression models and conditional estimates.

Independent Variable	**(Model 1) General UWLS MRA**[a]	**(Model 2) Specific UWLS MRA**	**(Model 3) Specific FE PANEL MRA**
Constant (Model Intercept)	−0.006 (0.005)	−0.0005 (0.0004)	−0.0005* (0.0002)
SE	1.40* (0.16)	1.64* (0.28)	1.67* (0.07)
Published	−0.001 (0.002)	–	–
Annual	0.008* (0.003)	–	–
Full Marketing Mix	0.003 (0.006)	–	–
Abs: GRP	0.02 (0.009)	0.03* (0.009)	0.01* (0.005)
Competition	0.002 (0.004)	–	–
Growth	0.002* (0.009)	–	–
Omit: Lag Price	0.009* (0.005)	–	–
Entertainment Media	0.07* (0.01)	0.07* (0.007)	0.01 (0.02)
Panel	0.002 (0.002)	–	–
US	0.006* (0.003)	0.007 (0.18)	0.004* (0.001)
TV	0.006* (0.003)	–	–
Internet	0.003* (0.008)	–	–
Category	−0.01* (0.005)	−0.005* (0.002)	−0.003 (0.002)
Display	−0.05 (0.03)	–	–
Dynamic	0.002 (0.003)	–	–
Dep: Market Share	−0.001 (0.004)	–	–
Abs: Monetary Spending	−0.008* (0.004)	–	–
Omit: Lag Advertising	0.002 (0.003)	–	–
Durables	0.001 (0.009)	–	–
Long-term	0.007 (0.006)	–	–
Endogeneity	0.01 (0.007)	–	–
Time Trend	−0.0005 (0.003)	–	–
Dep: Choice	−0.0009 (0.002)	–	–
Seasonality	−0.0006 (0.002)	–	–
Recession	0.05* (0.02)	0.05* (0.004)	0.006* (0.001)
Conditional Estimate Range[b]	0[.05*, 0.10*]	0[.04*, 0.05*]	0[.003, 0.02*]

[a]We pluralistically look across models. See Web Appendix F and this work for more details on effect estimation.
[b]We estimate numerous conditional estimates; context-specific effects vary more than our conditional range. See Web Appendix F and our included code for details of how we estimate both.
Notes. Numbers in parentheses are cluster-robust SE. * indicates $p < 0.05$. n = 1096 estimates. S = 93 studies.

which is indicative of, but does not prove, publication selection bias. We provide study-level boxplots and descriptive statistics on both the complete sample of coded estimates as well as on the final database of usable estimates used in Web Appendix H. We note that the descriptive statistics for all variables in both cases are qualitatively similar, suggesting no evidence of a significant difference across samples. Also, elasticities become small as the precision increases (Figs. 1-2).

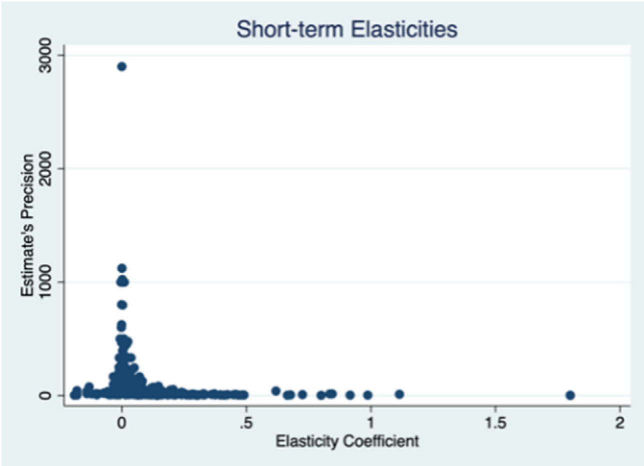

Fig. 1. Funnel plots: short and long-term effects.

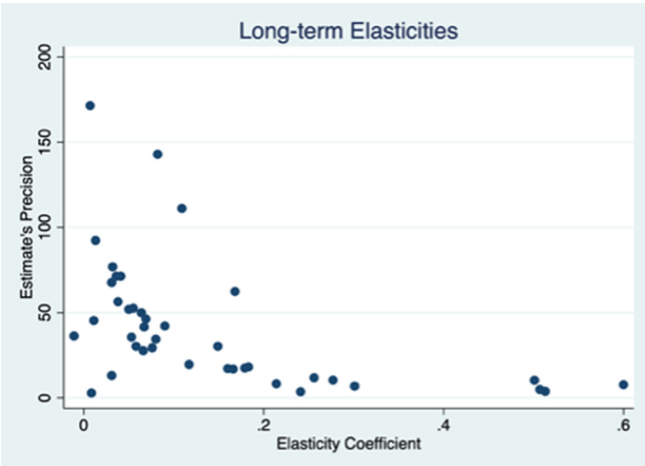

Fig. 2. Funnel plots: short and long-term effects.

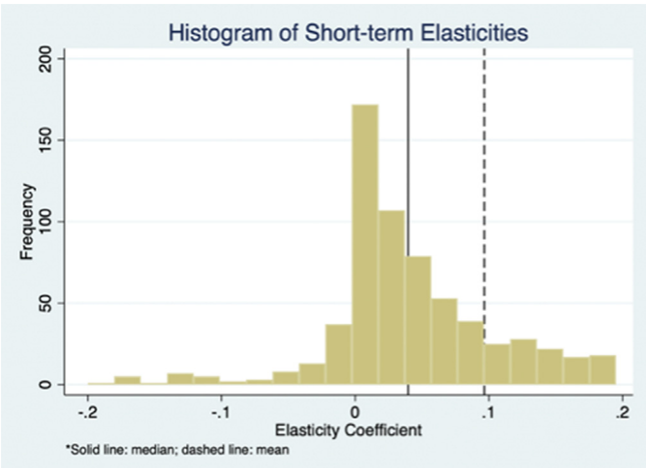

Fig. 3. Histograms: short-term and long-term elasticity effect sizes.

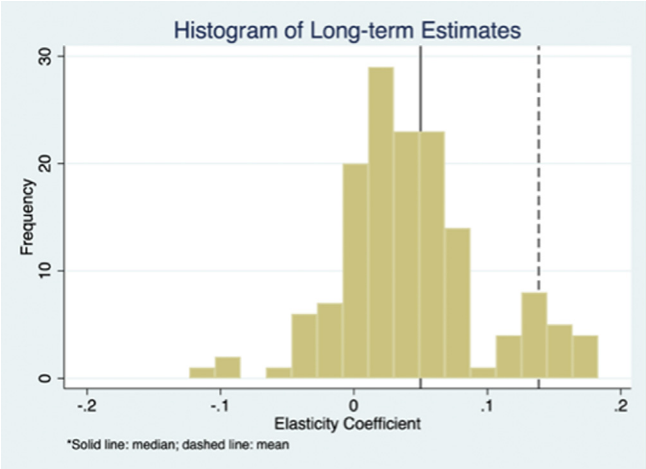

Fig. 4. Histograms: short-term and long-term elasticity effect sizes.

4.2. Meta-Analysis (MA) weighted averages

Next, we report overall simple averages, MA weighted averages, and MRA overall effect estimates with and without publication selection bias tests and corrections (Table 1 and Table WI4), the first 4 steps of our 8-step approach. Table 1, Model 1 allows us to compare our data and findings to results one obtains when applying the *meta*-analyses methods employed by the prior literature-descriptive averages (Sethuraman et al., 2011). Note that steps 1 – 8 of our method and these progressively more sophisticated models, as we move from left to right in both Tables 1-2, are 8 different competing methods; we replicate prior research in Model 1 of Table 1. Models 2–4 in Table 1 are more sophisticated but still inferior methods. Our preferred models examine conditional effects (Table 2) as opposed to overall effects as past research has done (Table 1).

In Model 1 of Table 1, we first report the simple unadjusted average short-term and long-term elasticities in Model 1 (Step 1), which replicates the overall effects reported in Sethuraman et al. (2011) and Henningsen et al. (2011) on our database.[1] These simple averages are very close to the averages (0.09 for short-term and 0.19 for long-term) reported by Henningsen et al. (2011) because our coding and inclusion rules are closest to theirs.

To adjust for known sampling error variances and to correct for heterogeneity, we next report RE MA in Table 1 (Model 2) following Equation (1) (Step 2). Although both RE estimates are statistically significant, they are much smaller than the corresponding simple averages, by over 70 % for short-term and by nearly 50 % for long-term elasticities. Furthermore τ^2 is small, suggesting little heterogeneity; $\tau = 0.014$ for the short-term estimates and somewhat more ($\tau = 0.045$) for the long-term estimates.

We then report UWLS, following Equation (2) in Table 1, Model 3 (Step 3). UWLS' short-term estimate is 0.004, which is very small, yet statistically significant. This estimate implies that a 10 % increase in advertising is associated with a .04 % increase in sales, market share, or choice. In the long-term, our UWLS estimate suggests a 10 % increase in advertising is associated with a 0.5% increase in sales.

In Model 4 of Table 1 following Equation (3) (Step 4), to further correct for potential publication selection bias via Funnel Asymmetry Test – Precision Effect Test Meta Regression Analysis (FAT-PET MRA), we report the preferred MA average of 0.0008 (short-term) and 0.03 (long-term), which is much smaller still than UWLS.

Methodological improvements are the main reason for this deflation of the estimated advertising elasticity given the application to the same data set. This central finding holds across all subgroups and models reported in Web Appendices H – J.

Via results in Table 1, Model 4, we offer clear evidence of funnel asymmetry, consistent with publication selection bias, befitting the visual inspection on funnel asymmetry above (see Figs. 1 and 2). When we account for publication selection bias, we find even smaller elasticity estimates, as in Model 4 compared to Model 3. Also, TESS-PSST is highly significant ($p < 0.001$) in all subgroups examined in Web Appendix J, corroborating the prevalent PSB we find in the literature via FAT-PET MRA. Via 3PSM, we find evidence of PSB in 3 of the 5 subgroups examined (one subgroup did not converge). We provide more details in Web Appendices H and J.

[1] For a summary of the coding differences between Sethuraman et al. (2011) and Henningsen et al. (2011), see Henningsen et al. (2011, p. 197 – 198). For the studies we include in our database, see Figures WG1 – WG2.

In Web Appendix H, we also report salient subgroup (on sales, market share and choice as dependent variable) MA overall effect estimates, to exclude potential aggregation bias as a logic driving our results. Via Tables WH1 – WH3, we show that sales and market share estimates are qualitatively similar to each other and those in Table 1, but estimates based on choice outcomes are smaller still.

In sum, these results identify our first major finding: via all MA weighted averages methods, we estimate overall advertising effects to be notably smaller than previous *meta*-analyses.[2] Furthermore, the size of SE's coefficient is quite substantial (1.90 and 1.86). To understand how large this is, suppose, hypothetically, that the mean elasticity is 0 and every estimate is reported as statistically significant and positive. Then, SE's coefficient is only somewhat larger, 2.34, as 2.34 is the mean of a standard normal distribution truncated at 1.96.

4.3. Multiple meta-regression analysis: explaining research variation and correcting biases

Next, we use MMRA to detect and filter out potential biases and explain differences among reported elasticities (i.e., heterogeneity), following MAER-Net guidelines (Havránek et al. 2020).

4.3.1. General MMRA model

We code 25 firm, data, and method factors that characterize the primary studies (for an overview see Tables 2-4) based on those factors coded in prior advertising elasticity *meta*-analyses, those present in the primary studies examined, and the MA methods literature. As we show in both Tables 1 and 2, and Web Appendix J, SE is always identified, statistically, as the main explanatory factor. Therefore, to be careful and conservative, we set SE to zero for our conditional estimates. We show the association of each variable to the various bias categories and highlight which value should be set for best practices bias alleviation via Table 3 (Stanley & Doucouliagos, 2012). If the academic publication process is, in part, responsible for publication selection bias, then unpublished papers (Published = 0) may be less biased. However, we note the opposite, suggesting this bias may be present prior to publication, as has been recently shown to often be the case in economics (Brodeur et al., 2023; Chopra et al., 2023).

We first estimate the general MMRA (as in Equation (4), above), including all 25 factors detailed in Table 3 as covariates, and report our findings in Model 1 of Table 2 (Step 5). The estimated coefficients from this model, as well as the methodology we describe here, can be used by industry analysts to estimate context-specific effects that apply more narrowly to their needs (e.g., advertising effects expected for specific input and output measure combinations in a specific region, industry, and chosen advertising platform(s)). We include more details on our conditional effect range and estimating context-specific effects in Web Appendix G.

A key variable of interest is *SE*, of which the effect is positive and significant, again pointing at publication selection bias – studies with higher SE find higher elasticities, on average, as also observed, model-free, in Figs. 1 and 2. The positive and significant covariate *Annual* shows estimates using annual data to be larger than those based on higher frequency data, consistent with the concept and prior findings that when less frequent data intervals are used, the portion of advertising identified as the current period is longer (Sethuraman et al., 2011). The coefficient of *Full Marketing Mix* shows models that account for a wide variety of marketing mix instruments find smaller effects, showcasing models with omitted marketing instruments may report upward biased elasticities (i.e., endogeneity). We find elasticities estimated on products in the *Growth* phase (versus maturity phase of the life cycle) to be larger, in line with the informative role of advertising (Narayanan, Manchanda, & Chintagunta, 2005). The coefficient of *Omit: Lag* shows that estimates that do not control for past prices find larger advertising effects. Advertising elasticities in studies in the US are larger than outside the US. The positive coefficient of *Entertainment Media* shows that advertising is more effective in the entertainment industry, consistent with the informative role of advertising in this area and recent findings (Schöndeling et al., 2023; Vakratsas, Demetrios & Ambler, 1999). The positive coefficients on TV and Internet show that these media are more effective, with Internet-based advertising being most effective.

4.3.2. Specific MRA models

To address possible multicollinearity among the 25 factors we include, we reduce the general model in Model 1 of Table 2 to those research dimensions that are statistically most important (see Model 2 of Table 2, based on Equation (4)) (Step 6). Note we perform a more sophisticated variable selection with Bayesian model averaging (BMA) in Web Appendix I (it shows similar results to those we report here). The results of the specific models are similar to the general model. In this more compact model, four coefficients now gain significance ($p < 0.05$).

First, *published* studies show slightly smaller elasticities than non-published papers. Secondly, estimates measured in gross ratings points (*Abs: GRP*) are typically larger than when other measures are used; this is a logical result and confirms past findings as a 1 % change in advertising dollars spend yields a less than 1 % change in GRP acquired (Sethuraman et al., 2011). Thirdly, accounting for *competition* results in smaller effects, likely again because of endogeneity from omitted vari-

[2] We do not claim that this finding is yet fully established. Here, we merely signal the emergence of one of our core messages that we strengthen through the many additional analyses and robustness checks that follow.

ables overexpressing the elasticity. Fourthly, studies that use panel data structures find smaller effects, possibly because panel data are typically less sensitive to endogeneity concerns (Ebbes, Papies, & van Heerde, 2021).

In Table 2, we also report Model 3 by additionally applying fixed effects for studies, following Equation (5) (Step 7). This represents our preferred model. Note fewer variables remain statistically significant once we further correct for study-level dependencies via fixed effects panel.

As one can see from Table 2, these models again provide very similar estimates. For all models, we also provide a range of 'conditional' estimates in Appendix F by substituting specific values into the independent variables, as summarized by the final row of Table 2 (Step 8). These ranges encompass the potential the short and long-term effects expected across the various generalizable scenarios examined. They account for PSB and best research practices to reduce misspecification bias, while allowing aggregation-related variables to fluctuate, thus establishing a range of possible effects. Only in very specific and favorable scenarios do effects begin to come close to the overall average effect estimates reported by past literature (such as internet-based advertising in the entertainment industry for growth-stage products in the US estimated using annual data). We refer to Web Appendix K for more explanation on terminology and abbreviations.

As we show in Table 3, variables associated with aggregation bias can be seen as potential conditional heterogeneity of effect from aggregating different scenarios. The value that should be set for most aggregation-related variables depends on whether the user is estimating generalizable conditional estimates, as we are here, or if they are instead estimating a specific context that interests the user. That is, aggregation bias can be alleviated by either setting values for more generalizable effects or by focusing the advertising effectiveness estimate on the context of interest, alongside various attributes ranging from how the input (advertising) or output (sales, market share, or choice) is measured, to what advertising medium is used, to the industry or region in which the product or firm operates, the stage in the product life cycle, and beyond. This is done via conditional effect estimates, such as those shown in the final row of Table 2. We have reproduced estimates similar to past research under conditions consistent with their approach (Table 1, Model 1) and then progressively layered on citation-supported corrections throughout the rest of Tables 1 – 2.

4.4. Robustness checks

In this section, we summarize numerous robustness checks including alternative PSB testing and estimation methods, trend analysis, and subgroup analyses; all corroborate our findings.

4.4.1. Alternative methods to accommodate publication selection bias

Utilizing TESS-PPST, we find that both TESS and PSST separately and combined detect publication selection bias in the aggregate database of elasticities and in every subgroup examined, as we show in Web Appendix K, Table WK1. This further verifies our finding that there is significant publication selection bias in the advertising effectiveness literature.

We also include and describe three alternative *meta*-analytic weighted averages used to accommodate publication selection biases by calculating UWLS weighted averages of those estimates with statistical power > = 80 % for aggregate, short-term, and long-term effects in Web Appendix J, Table WJ2. The aggregate and short-term effects are both estimated as being 0.02 and not statistically different than zero ($p > 0.05$ in both cases). Only the long-term effect is significantly positive (0.05, $t = 3.81$; $p < 0.01$). These estimates further support our finding that overall advertising elasticity estimates are much smaller than prior *meta*-analytic estimates.

4.4.2. Time trends in advertising elasticity as alternative explanation

It may be that our estimates are smaller due to the possibility that elasticities have potentially decreased over time; however, this is not the case. We show that, over time, the reported MA estimates of advertising-sales elasticities have become smaller, largely due to the advancement in MA methods. Like our study, Sethuraman et al. (2011) also finds much smaller advertising elasticity than prior *meta*-analyses: 0.12 versus. 0.22 for the short-term and 0.24 versus 0.41 for the long-term (p. 457). We find a similar decline from Sethuraman et al. (2011). However, this decrease is not caused by a decline of the elasticities reported in the primary research literature. If anything, they have been increasing. We estimate an increase in elasticity of 0.00004/year, implying average elasticity estimates may have increased by 0.02 over the 50-year history of this research.

4.4.3. Small sample as alternative explanation: expanding the sample

In this work, we follow past studies focusing on brand-level elasticities. However, to expand our sample and further demonstrate the robustness of our findings, we additionally estimate models similar to those of Tables 1 and 2 in Tables 3-4, but using a much larger sample of elasticities. In Tables 3 - 4, we consolidate brand and category elasticities using 1,096 estimates from 93 papers. We find similar estimates from inferior models as we do in Table 1, and similar reductions and corrected estimates as we progress from worse to better models across Tables 3 - 4, again, similar to that which we find in Tables 1 – 2. We find, in the wider sample, the database attributes measured by a variety of descriptive distributional statistic for all estimates, as opposed to the smaller subset with SE, are similar to each other and to those of the final sample used in our main analyses in Tables 1 – 2.

By allowing for a common regression between brand and category elasticities using the initial variable set, again allowing general-to-specific to select our model, and then applying conditional estimate analyses to the models to estimate a range of typical conditional effects, we estimate the effect range applicable for the majority of scenarios. However, it should be noted that these effects are only robust if we assume no interaction effects and similar relations between the variables included in the model for both brand and category elasticities (unlikely). These estimates suggest that elasticities, generally speaking, are even smaller, and that category elasticities in particular are smaller but more compact, falling into a thinner typical range of 0.003 – 0.10 except in the entertainment media industry where effects are larger (~ 0.07 – 0.10 larger across models).

We consider the FAT-PET MMRA model with fixed effects in Column (3) of Table 5 the best practice method and, thus, consider these estimates as approximating the "true" estimates most closely. However, we prefer Tables 1 – 2 to Tables 4 – 5 for estimating brand-level effects. Tables 4 – 5 are estimated merely for robustness.

4.4.4. Subgroup MMRA

First, we calculate the UWLS-MMRA model from Table 2, Model 2, using only those estimates with precision less than 200 to be sure that a few estimates with relatively large precisions (1/SE) are not unduly influencing the findings. The results are practically equivalent to those we report in Table 2, Model 2 (see Table WJ3, Model 1 in Web Appendix J). Secondly, to ensure our results are not influenced by any single coding error or other mistake, we investigate leverage statistics (dfbeta) and re-estimate the UWLS-MMRA model reported in Table 2, Model 2 (see Table WJ3, Model 2). We find qualitatively similar effects. Next, we again calculate the UWLS-MMRA but using only those studies from Henningsen et al. (2011) (see Table WJ3, Model 3). The results are similar but not identical. For example, effects from TV-based advertisements exhibit a notably smaller, rather than larger, effect. We hypothesize this is because the Internet variable must be omitted due to collinearity. In Table WJ3, Model 4, we replicate Table 2, Model 3, employing a random effects panel instead of fixed effects panel.

4.4.5. Conclusions additional analyses

These robustness checks corroborate the significant reductions in advertising elasticity that we find are due to improved methods. In summary, we apply multiple *meta*-analytic models to various models and subgroups and find an effect range that is robust to subgroup analyses, alternative estimation methods, and multiple tests for and methods to accommodate publication selection bias.

5. Discussion

In this section, we discuss the size and impact of the biases that we identified and corrected for in the previous section in the advertising effectiveness context; implications for scholars and practice, limitations of our research, and future research directions.

5.1. Meta-analytic bias

Prior *meta*-analyses reported notably larger overall advertising effects because they did not fully adjust for common *meta*-analytic biases. We detail the sources and impact each of these biases in *meta*-analytic advertising research specifically. The mathematical magnitude of each bias and its effect on the elasticity estimates can be inferred by the coefficients on the independent variables in those categories.[3]

5.1.1. Publication selection bias

First and foremost, we find strong and robust evidence that reported advertising elasticities are overestimated due to publication selection bias. We demonstrate this by the large size of the coefficient on the SE and its consistent statistical significance ($p < 0.01$) across all methods, models, and subgroup analyses. Whether we use simple *meta*-analytic weighted averages and methods (Table 1) or any MMRA model (Table 2), SE's coefficient is always quite substantial (1.42–1.75). To see what these large coefficients imply, take the smallest (1.42) and multiply by the average standard error (0.0532), yielding 0.08. Thus, publication selection bias exaggerates reported advertising elasticities by at least 0.08, on average. Subtracting average bias from the average reported advertising elasticity (0.09) implies that the typical publication selection bias-corrected elasticity is less than 0.02.

Overall, we find that the average reported advertising elasticity is exaggerated by a factor of 5.7, or nearly six times for long-term estimates and even more for short-term estimates. Suppose, for the sake of argument, that SE does not represent publication selection bias. A conscientious researcher would still drive standard error to zero. Bias or not, standard error represents sampling error and the uncertainty of the associated elasticity estimates.

Nonetheless, this exaggeration is likely to be the result of selection for statistical significance, as our tests have indicated. That is, some researchers and/or reviewers suppress negative and/or nonsignificant elasticities, potentially believing them to

[3] For example, one aspect of misspecification bias is represented in our data by the variable "Panel". Since we define 'Panel = 1' to be when (superior) panel data was used and since the size of the coefficient on that variable is 0.001 in our preferred method (Table 2, Model 3), we can say that we estimate the misspecification bias associated with the use of inferior data structures leads to underestimating elasticity effects by 0.001. We can sum the relevant coefficients in this category appropriately, based on best practice values, to estimate the overall effect of misspecification on elasticity estimates formed by the overall *meta*-analytic average or other averages.

be mistaken, or the models and methods that produced them to be at fault. After all, the idea that advertising increases sales is a core principle of marketing. This asymmetry of theory leads to the 'fallacy of composition.'

5.1.2. Misspecification bias

We have also identified several other potential sources of bias. Notable misspecification biases are represented by Panel, Competition, and Omit: Lag Price, Model 3, Table 2. We find that studies that do not control for the panel structure of advertising data or changes in product price tend to underestimate effects, while studies that do not control for competition tend to overestimate effects. However, these potential misspecification biases are quite small (< ± 1 %).

5.1.3. Aggregation bias

Aggregation bias can also be notable; however, it is only a 'bias' when the estimate is misaligned with its target industry, region, advertising platform, or the input or output measure used. We coded and controlled for a multitude of factors representing industries, outcome measures, advertising mediums, and sales measures in the 'general' MMRA model (Table 2, Model 1). The potential aggregation bias that might possibly remain in the 'specific' MMRA (Table 2, Model 2) is associated with the following factors: Entertainment Media, Abs: GRP, Growth, US, Internet, Annual, and TV. These aggregation biases can be substantial. If a *meta*-analytic study fails to account for the different responsiveness of advertising in the entertainment media industry (or those in their growth phase), advertising effectiveness estimates will be underestimated. However, in other industries and those in their mature stage, these will be overestimated. Similarly, if combined indiscriminately, advertising is likely to be underestimated for TV and Internet platforms, and overestimated for others.

When all of these considerations are taken into account, representative advertising elasticities are quite modest except in particular scenarios; for example, US-based entertainment media firms advertising growth phase products (e.g., tickets to new movies) using internet advertising. Overall effect estimates from *meta*-analytic weighted averages and those that account for publication selection bias are often practically negligible (see Table 2).

5.2. Implications

This research brings forth critical implications for both scholarly pursuits and marketing practice, significantly extending the discourse on advertising effectiveness. We underscore the need for a paradigm shift in how advertising effectiveness is assessed and used in strategic decision-making.

5.2.1. Implications for scholars and analysts

Our results suggest important variables and modeling considerations for (1) primary studies in advertising; (2) MA of advertising; and (3) marketing MA generally.

First, primary study advertising researchers should incorporate the factors we find relevant here in their models. Specifically, advertising modelers should incorporate the lag of price in their advertising effectiveness models, account for different advertising mediums used in the data they are modeling, control for changes to other marketing mix elements, account for competition, and use panel data when possible. Analysts and readers alike should also be cognizant that studies that examine advertising over longer periods and use annual data are likely to find larger effects.

Secondly, this work empirically demonstrates for future *meta*-analysts in advertising a minimum set of variables that modelers should account for when replicating and/or updating our findings. Better methods reduce the number of significant variables extensively, compared to prior findings, offering future modelers guidance in terms of which variables are most important.

Thirdly, this work provides evidence as to why marketing *meta*-analysts generally should incorporate cluster-robust standard errors, employ UWLS, include the standard error as an independent variable, and use study-level fixed effects when estimating *meta*-analytic models, as well as the impact of doing so. Marketing *meta*-analysts can use the best practice bias-correction methods we demonstrate using the 8-step approach advocated here, thereby providing a future path for *meta*-analysts to follow.

Fourthly, this research should spark further demand for increased scrutiny in marketing research reporting. The significant publication selection bias uncovered by our advanced *meta*-analytic methods indicates that previous estimates of marketing effectiveness are likely inflated. This discovery mandates a more rigorous evaluation approach. By implementing our proposed methodologies, analysts and firms can ensure a more accurate, honest, and transparent evaluation of marketing efforts, facilitating better-informed strategic decisions.

5.2.2. Implications for practice

Our research findings compel a strategic reassessment of marketing budgets and resource allocation. With our findings indicating a more nuanced view of advertising effectiveness, firms should consider reallocating budgets away from traditional advertising to potentially more lucrative areas, such as digital marketing and customer relationship management. These areas often offer more measurable and direct returns, particularly in segments showing higher responsiveness.

Our findings are particularly valuable for firms operating under economic strain or looking to optimize budgetary allocations. By demonstrating variable advertising effectiveness, our study guides companies in making judicious budget cuts without sacrificing impact. Firms are encouraged to invest in high-return areas while reducing spending in less effective ones, thus maximizing overall marketing efficiency.

We provide an overall mean and relevant range that can be used for benchmarking. Marketing managers and analysts can use these results to better allocate their advertising / marketing budgets, as well as to evaluate the quality of advertising campaigns. In general, practitioners can expect a 10 % increase in advertising to be associated with a 0.03 to 1 % increase in sales, choice, and/or market share.

Furthermore, we also show when advertising is most effective. Our results show that advertising can have an even larger effect in particular contexts, such as in the entertainment media industry, when internet advertising is employed, and/or during the growth stage. Thus, firms receive more value for money when focusing their advertising efforts on the internet, primarily, and TV. Advertising is most effective during the growth phase of the product life cycle. These results also should encourage marketing departments to enhance their accountability by adopting advanced analytics. This shift towards data-driven decision-making ensures that advertising spending is justified by measurable outcomes, improving overall marketing strategy effectiveness.

Additionally, managers could benefit from optimizing advertising strategies based on our findings. Recognizing the varied effectiveness across different mediums and stages of the product life cycle, strategic targeting of advertising efforts can lead to significantly improved returns. For instance, focusing on internet and TV advertising during the growth stage can maximize impact.

Practitioners can also apply our methodologies to tailor their advertising strategies to specific conditions. By using the conditional estimation models provided, marketing professionals can adjust their strategies according to industry-specific dynamics, regional characteristics, and other relevant factors, enhancing the precision and effectiveness of their advertising efforts. We provide estimated coefficients, such as those from our general model in Table 2, Model 1, along with the associated methodologies that we describe here. Our data and code can be used by corporate or industry analysts to estimate specific effects that more narrowly apply to their needs.

Using conditional estimation based on a linear combination of the independent variables, practitioners can calibrate their projections based on our data and general model to fit their industry, region, chosen advertising platform, product life cycle stage, outcome of interest, and beyond (e.g., internet-based advertising sales effects for a 'mature' product in the entertainment media industry). To support *meta*-analysts, we provide all data and code to readers at www.MetaAnalysisSupport.com.

Lastly, our *meta*-analytic methods demonstration provides practitioners with their own *meta*-analytic how-to they can use to dig into and analyze other business and marketing issues using the techniques outlined here. Integrating our methodological advances allows firms to refine their marketing strategies based on robust empirical evidence. Our approach offers a detailed understanding of when and how advertising is most effective, enabling marketers to tailor strategies to specific conditions such as industry dynamics, regional characteristics, and product life cycle stages. Using our data and code as a guide, this work unlocks a new world for marketing practitioners to easily conduct meaningful research that provides robust market insights. With these methods, advertising effectiveness research need not be limited to RCTs. Our work shows that proper *meta*-analytic estimation of a collection of MMMs can provide accurate estimates of advertising elasticity in line with those of recent RCTs, broadening our ability to collect and conduct advertising effectiveness data and research under a wider range of settings and time periods.

5.3. Limitations and future research directions

First, our MA cannot predict what might happen if marketers dramatically alter their advertising efforts because such a dramatic shift is not typical of the much smaller sample variation upon which our estimates are based. Therefore, it would be erroneous to infer that advertising has no effect.

Competition and experience may have led to near optimal advertising levels at a mature stage of a product/brand's life cycle. When advertising and sales are in near equilibrium, marginal changes—the type of variation generally seen in the research studies we have coded—would be expected to have little effect. That is, if we assume efficient markets, firms are already at or near optimal advertising levels. In this case, the cost of additional advertising will be offset by additional benefits, providing no net gain. Furthermore, we estimate increases in both sales and market share associated with marginal increases in advertising spending. These two separate effects occur simultaneously; increases in sales and market share resulting from increasing advertising may produce a double benefit for the firm.

Secondly, we can only analyze 40 studies in our main analyses given we need standard errors, and many studies do not include standard errors or allow for them to be calculated. Due to the number of available estimates, we can only analyze up to 25 variables due to multicollinearity concerns. While one always prefers larger databases, the research on which can be fruitful, we feel the ability to weigh studies by their accuracy and account for PSB outweighs the benefits of coding a large number of estimates without being able to weight them or use advanced *meta*-analytic methods such as FAT-PET-PEESE or UWLS.

Thirdly, while this work examines and details best practice *meta*-analytic methods primarily for a frequentist approach, a budding and relevant field in MA is Bayesian MA and bias corrections. Marketing scholars should advance best practices in conducting *meta*-analyses in a Bayesian approach, in addition to the best practices suggested here for frequentist *meta*-analyses.

Fourthly, future *meta*-analyses in marketing should examine other research questions by applying these *meta*-analytic bias correction methods and should consider the 8-step approach we suggest here to quantify the bias and correct effect estimates in both new and other previously-studied areas. While advertising research has a large and still growing number of *meta*-analyses conducted, such areas benefitting from improved and updated *meta*-analyses concern price, cross-price, and promotional elasticities, among others, such as those areas listed in Korkames & Stremersch (2024) and Korkames (2024b).

CRediT authorship contribution statement

Joseph Korkames: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization, Formal analysis, Data curation. **T.D. Stanley:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis. **Stefan Stremersch:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Data availability

The data and code are available at mMetaAnalysisSupport.com.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijresmar.2025.04.001.

References

Alinaghi, N., & Reed, W. R. (2018). Meta-analysis and publication bias: How well does the FAT-PET-PEESE procedure work? *Research Synthesis Methods, 9*(2), 285–311. https://doi.org/10.1002/jrsm.1298.

Almalik, O. (2024). Adjusting for publication bias in meta-analysis with continuous outcomes: A comparative study. arXiv preprint arXiv:2410.06309. https://arxiv.org/pdf/2410.06309.

Assmus, G., Farley, J. U., & Lehmann, D. R. (1984). How advertising affects sales: Meta-analysis of econometric results. *Journal of Marketing, 21*(1), 65–74. https://doi.org/10.2307/3151793.

Bartoš, F., Maier, M., Wagenmakers, E.-J., Doucouliagos, H., & Stanley, T. D. (2022). No need to choose: Robust Bayesian meta-analysis with competing publication bias adjustment methods. *Research Synthesis Methods, 14*(1), 99–116. https://doi.org/10.1002/jrsm.1594.

Batra, R., Myers, J. G., & Aaker, D. A. (1996). *Advertising management*. Prentice Hall.

Bijmolt, T. H. A., & Pieters, R. G. M. (2001). Meta-analysis in marketing when studies contain multiple measurements. *Marketing Letters, 12*(2), 157–169.

Bom, P. R., & Rachinger, H. (2019). A kinked meta-regression model for publication bias correction. *Research Synthesis Methods, 10*(4), 497–514. https://doi.org/10.1002/jrsm.1352.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97–111. https://doi.org/10.1002/jrsm.12.

Borenstein, M., Rothstein, H. R., Hedges, L. V., & Higgins, J. (2009). Introduction to meta-analysis. *John Wiley and Sons*. https://doi.org/10.1002/9780470743386.

Borenstein, M., Hedges, L., & Rothstein, H. (2007). Meta-analysis fixed effect vs. random effects. www.meta-analysis.com. https://meta-analysis.com/download/M-a_f_e_v_r_e_sv.pdf.

Brodeur, A., Carrell, S., Figlio, D., & Lusher, L. (2023). Unpacking p- hacking and publication bias. *American Economic Review, 113*(11), 2974–3002.

Campos, J., Ericsson, N. R., & Hendry, D. F. (2005). General-to-specific modeling: An overview and selected bibliography. *International Finance Discussion Paper, 2005*(838), 1–94. https://doi.org/10.17016/ifdp.2005.838.

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science, 2*(2), 115–144. https://doi.org/10.1177/2515245919847196.

Chopra, F., Haaland, I., Roth, C., & Stegmann, A. (2023). The null result penalty. *Economic Journal early online.*.

Clarke, D. G. (1976). Econometric measurement of the duration of advertising effect on sales. *Journal of Marketing Research, 13*(4), 345. https://doi.org/10.2307/3151017.

Danaher, P. J., & Brodie, R. J. (2000). Understanding the characteristics of price elasticities for frequently purchased packaged goods. *Journal of Marketing Management, 16*(8), 917–936. https://doi.org/10.1362/026725700784683726.

Ebbes, P., Papies, D., & van Heerde, H. J. (2021). Dealing with endogeneity: A nontechnical guide for marketing researchers. In *Handbook of Market Research* (pp. 181–217). Springer.

Edeling, A., & Fischer, M. (2016). Marketing's impact on firm value: Generalizations from a meta-analysis. *Journal of Marketing Research, 53*(4), 515–534. https://doi.org/10.1509/jmr.14.0046.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629.

Eisend, M. (2009). A meta-analysis of humor in advertising. *Journal of the Academy of Marketing Science, 37*(2), 191–203. https://doi.org/10.1007/s11747-008-0096-y.

Eisend, M. (2015). Have we progressed marketing knowledge? A meta-meta-analysis of effect sizes in marketing research. *Journal of Marketing, 79*(3), 23–40. https://doi.org/10.1509/jm.14.0288.

Eisend, M., & Tarrahi, F. (2014). Meta-analysis selection bias in marketing research. *International Journal of Research in Marketing, 31*(3), 317–326. https://doi.org/10.1016/j.ijresmar.2014.03.006.

Eysenck, H. J. (1984). Meta-analysis: an abuse of research integration. *The Journal of Special Education, 18*(1), 41–59. https://doi.org/10.1177/002246698401800106.

Farley, J. U., Lehmann, D. R., & Sawyer, A. (1995). Empirical marketing generalization using meta-analysis. Marketing Science, 14(3, Part 2), G36–G46. doi: 10.2307/184146.

Gelman, Andrew (2018), The P-curve, P-uniform, and Hedges (1984) methods for meta-analysis under selection bias: an exchange with Blake McShane, Uri Simonsohn, and Marcel van Assen. Retrieved February 10, 2021, from https://statmodeling.stat.columbia.edu/2018/02/26/p-curve-p-uniform-hedges-1984-methods-meta-analysis-selection-bias-exchange-blake-mcshane-uri-simosohn/.

Gordon, B. R., Taddy, M., & Zettelmeyer, F. (2022). Close enough? A large-scale exploration of non-experimental approaches to advertising measurement. *International Journal of Research in Marketing, forthcoming* https://arxiv.org/abs/2201.07055.

Gordon, B. R., Zettelmeyer, F., Bhargava, N., & Chapsky, D. (2019). A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook. *Marketing Science, 38*(2), 193–225. https://doi.org/10.1287/mksc.2018.1135.

Greene, W. H. (1990). Econometric analysis. *Macmillan*. https://doi.org/10.1017/s0266466600004278.

Grewal, D., Puccinelli, N., & Monroe, K. B. (2017). Meta-analysis: Integrating accumulated knowledge. *Journal of the Academy of Marketing Science, 46*(1), 9–30. https://doi.org/10.1007/s11747-017-0570-5.

Gujarati, D. N. (1995). *Basic econometrics*. McGraw-Hill.

Hanssens, D. M. (2015). *Empirical generalizations about marketing impact*. Marketing Science Institute.

Hanssens, D. M. (2018). The value of empirical generalizations in marketing. *Journal of the Academy of Marketing Science, 46*(1), 6–8. https://doi.org/10.1007/s11747-017-0567-0.

Havránek, T., Stanley, T. D., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W. R., Rost, K., & van Aert, R. C. M. (2020). Reporting guidelines for meta-analysis in economics. *Journal of Economic Surveys, 34*(3), 469–475. https://doi.org/10.1111/joes.12363.

Havránek, T., Irsova, Z., Laslopova, L., & Zeynalova, O. (2022). Publication and attenuation biases in measuring skill substitution. *Review of Economics and Statistics, 1–72*. https://doi.org/10.1162/rest_a_01227.

Henmi, M., & Copas, J. B. (2010). Confidence intervals for random effects meta-analysis and robustness to publication bias. *Statistics in Medicine, 29*, 2969–2983. https://doi.org/10.1002/sim.4029.

Henningsen, S., Heuke, R., & Clement, M. (2011). Determinants of advertising effectiveness: The development of an international advertising elasticity database and a meta-analysis. *Business Research, 4*(2), 193–239. https://doi.org/10.1007/bf03342755.

Horsky, D., Horsky, S., & Zeithammer, R. (2016). The modern advertising agency selection contest: A case for stipends to new participants. *Journal of Marketing Research, 53*(5), 773–789. https://doi.org/10.1509/jmr.14.0347.

IntHout, J., Ioannidis, J. P. A., Borm, G. F., & Goeman, J. J. (2015). Small studies are more heterogeneous than large ones: A meta-meta-analysis. *Journal of Clinical Epidemiology, 68*, 860–869. https://doi.org/10.1016/j.jclinepi.2015.03.017.

Iyengar, S., & Greenhouse, J. B. (1988). Selection models and the file drawer problem. *Statistical Science, 3*(1), 109–117.

Kolsarici, C., Vakratsas, D., & Naik, P. A. (2020). The anatomy of the advertising budget decision: How analytics and heuristics drive sales performance. *Journal of Marketing Research, 57*(3), 457–478. https://doi.org/10.1177/0022243720907578.

Korkames, J. (2024a). Chapter 1: A comprehensive guide to meta-analysis and meta-regression in marketing research: Methods, applications, and best practices. In Estimating marketing's power via improved meta-analysis: A deep dive into advertising effectiveness (Doctoral dissertation, IESE Business School, University of Navarra). Retrieved from https://dadun.unav.edu/entities/publication/f82f9d39-2356-4fcc-8629-6837c20c595d

Korkames, J. (2024b). Price elasticity: Practical insights. In A. Hinterhuber (Ed.), Elgar encyclopedia of pricing (pp. 176–184). Edward Elgar Publishing. doi: 10.4337/9781035307319.ch31.

Korkames, J., & Stremersch, S. (2024). Price elasticity, cross-price elasticity, and advertising elasticity: Concepts and summary of prior meta-analytic insights. In A. Hinterhuber (Ed.), *Elgar Encyclopedia of Pricing* (pp. 168–175). Edward Elgar Publishing.

Köhler, C., Mantrala, M. K., Albers, S., & Kanuri, V. K. (2017). A meta-analysis of marketing communication carryover effects. *Journal of Marketing Research, 54*(6), 990–1008. https://doi.org/10.1509/jmr.13.0580.

Lavidge, R. J., & Steiner, G. A. (1961). A model for predictive measurements of advertising effectiveness. *Journal of Marketing, 25*(6), 59–62. https://doi.org/10.1177/002224296102500611.

Lewis, R. A., & Reiley, D. H. (2014). Online ads and offline sales: Measuring the effect of retail advertising via a controlled experiment on Yahoo! *Quantitative Marketing and Economics, 12*(3), 235–266. https://doi.org/10.1007/s11129-014-9146-6.

Lin, L., Shi, L., Chu, H., & Murad, M. H. (2020). The magnitude of small-study effects in the Cochrane database of systematic reviews: An empirical study of nearly 30,000 meta-analyses. *BMJ Evid Based Med, 25*(1), 27–32. https://doi.org/10.1136/bmjebm-2019-111191.

Lodish, L. M., Abraham, M. M., Livelsberger, J., Lubetkin, B., Richardson, B., & Stevens, M. E. (1995). A summary of fifty-five in-market experimental estimates of the long-term effect of TV advertising. *Marketing Science, 14*(3), G113–G140. https://doi.org/10.1287/mksc.14.3.g133.

MacKenzie, S. B., & Lutz, R. J. (1989). An empirical examination of the structural antecedents of attitude toward the ad in an advertising pretesting context. *Journal of Marketing, 53*(2), 48–65. https://doi.org/10.1177/002224298905300204.

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: an evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science, 11*(5), 730–749. https://doi.org/10.1177/1745691616662243.

Mawdsley, D., Higgins, J. P. T., Sutton, A. J., & Abrams, K. R. (2017). Accounting for heterogeneity in meta-analysis using a multiplicative model-an empirical study. *Research Synthesis Methods, 8*(1), 43–52. https://doi.org/10.1002/jrsm.1216.

Moreno, S. G., Sutton, A. J., Ades, A. E., Stanley, T. D., Abrams, K. R., Peters, J. L., & Cooper, N. J. (2009). Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Medical Research Methodology, 9*(1), 1–17. https://doi.org/10.1186/1471-2288-9-2.

Naik, P. A., & Raman, K. (2003). Understanding the impact of synergy in multimedia communications. *Journal of Marketing Research, 40*(4), 375–388. https://doi.org/10.1509/jmkr.40.4.375.19386.

Nakagawa, S., Lagisz, M., Jennions, M. D., Koricheva, J., Noble, D. W., Parker, T. H., Sánchez-Tójar, A., Yang, Y., & O'Dea, R. E. (2021). Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods in Ecology and Evolution, 13*(1), 4–21. https://doi.org/10.1111/2041-210X.13724.

Narayanan, S., Manchanda, P., & Chintagunta, P. K. (2005). Temporal differences in the role of marketing communication in new product categories. *Journal of Marketing Research, 42*(3). https://doi.org/10.1509/jmkr.2005.42.3.278.

Puto, C. P., & Wells, W. D. (1984). Informational and transformational advertising: The differential effects of time. *Advances in Consumer Research, 11*, 638–643.

Rockers, P. C., Røttingen, J.-A., Shemilt, I., Tugwell, P., & Bärnighausen, T. (2015). Inclusion of quasi-experimental studies in systematic reviews of health systems research. *Health Policy, 119*, 511–519. https://doi.org/10.1016/j.healthpol.2014.10.006.

Rosenberg, M. S. (2005). The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution, 59*(2), 464–468.

Rosengren, S., Eisend, M., Koslow, S., & Dahlen, M. (2020). A meta-analysis of when and how advertising creativity works. *Journal of Marketing, 84*(6), 39–56. https://doi.org/10.1177/0022242920929288.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Wiley.

Ruan, B., Hsee, C. K., & Lu, Z. Y. (2018). The teasing effect: an underappreciated benefit of creating and resolving an uncertainty. *Journal of Marketing Research, 55*(4), 556–570.

Rust, R. T., Lehmann, D. R., & Farley, J. U. (1990). Estimating publication bias in meta-analysis. *Journal of Marketing Research, 27*(2), 220. https://doi.org/10.2307/3172848.

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage Publishing.

Schöndeling, A., Burmester, A. B., Edeling, A., Marchand, A., & Clement, M. (2023). Marvelous advertising returns? A meta-analysis of advertising elasticities in the entertainment industry. *Journal of the Academy of Marketing Science.*. https://doi.org/10.1007/s11747-022-00916-0.

Sethuraman, R., Tellis, G. J., & Briesch, R. A. (2011). How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *Journal of Marketing Research, 48*(3), 457–471. https://doi.org/10.1509/jmkr.48.3.457.

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review, 17*(8), 881–901. https://doi.org/10.1016/s0272-7358(97)00056-1.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366. https://doi.org/10.1037/e519702015-014.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014a). P-curve: a key to the file-drawer. *Journal of Experimental Psychology: General, 143*(2), 534–547.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014b). P-curve and effect size: correcting for publication bias using only significant results. *Perspectives on Psychological Science, 9*(6), 666–681.

Stanley, T. D., & Doucouliagos, H. (2012). Meta-regression analysis in economics and business. *Routledge*. https://doi.org/10.4324/9780203111710.

Stanley, T. D., & Doucouliagos, H. (2015). Neither fixed nor random: Weighted least squares meta-analysis. *Statistics in Medicine, 34*(13), 2116–2127. https://doi.org/10.1002/sim.6481.

Stanley, T. D., Doucouliagos, H., Ioannidis, J. P. A., & Carter, E. C. (2021). Detecting publication selection bias through excess statistical significance. *Research Synthesis Methods, 12*(6), 776–795. https://doi.org/10.1002/jrsm.1512.

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance–Or vice versa. *Journal of the American Statistical Association, 54*(285), 30–34. https://doi.org/10.2307/2282137.

Vakratsas, D., & Ambler, T. (1999). How advertising works: What do we really know? *Journal of Marketing, 63*(1), 26–43. https://doi.org/10.2307/1251999.

van Aert, R. C. M., Wicherts, J. M., & van Assen, M. A. L. M. (2016). Conducting meta-analyses based on p values: Reservations and recommendations for applying P-Uniform and P-Curve. *Perspectives on Psychological Science, 11*(5), 713–729. https://doi.org/10.1177/1745691616650874.

van Heerde, J. V., Gijsenberg, M. J., Dekimpe, M. G., & Steenkamp, J.-B.-E.-M. (2013). Price and advertising effectiveness over the business cycle. *Journal of Marketing Research, 50*(2), 177–193. https://doi.org/10.1509/jmr.10.0414.

Yang, Y., Lagisz, M., Williams, C., Pan, J., Noble, D. W. A., & Nakagawa, S. (2023). Robust point and variance estimation for ecological and evolutionary meta-analyses with selective reporting and dependent effect sizes. *EcoEvoRxiv*. https://doi.org/10.32942/X20G6Q.