Economics
The Open-Access, Open-Assessment E-Journal

Discussion Paper
No. 2020-1 | January 07, 2020 | http://www.economics-ejournal.org/economics/discussionpapers/2020-1

Fixed-effect versus random-effects meta-analysis in economics: a study of pass-through rates for alcohol beverage excise taxes

Jon P. Nelson

Abstract
This paper compares two methods for meta-analysis: fixed-effect models and random-effects models. Both models are applied to pass-through rates of excise taxes on alcohol beverages. Using a sample of estimates from 30 primary studies, weighted means are first reported for each method and compared against a fully- passed tax or rate of unity. Dispersion and heterogeneity statistics are used to assess the performance of each method. Second, means and dispersion statistics are reported by subgroups for country source; beverage (beer, wine-spirits); and published status. Third, tests are conducted for publication selection bias using funnel plots and regression asymmetry tests. Fourth, three procedures are undertaken to reduce selection bias: trim-and-fill; cumulative meta-analysis; and meta-regressions. Based on a variety of tests and procedures, three conclusions are reached. First, a random-effects model is more appropriate for these data, reflecting diverse estimates of pass-through rates. Second, pass-through rates are approximately unity regardless of beverage. Third, greater attention needs to be given to choice of model for meta-analysis in economics.

JEL C18 C52 H21 H22 I18
Keywords Meta-analysis; fixed-effects; random-effects; publication bias; excise taxes; pass-through rate; alcohol

Authors
Jon P. Nelson, Pennsylvania State University, USA, jpn@psu.edu

Citation Jon P. Nelson (2020). Fixed-effect versus random-effects meta-analysis in economics: a study of pass-through rates for alcohol beverage excise taxes. Economics Discussion Papers, No 2020-1, Kiel Institute for the World Economy.
http://www.economics-ejournal.org/economics/discussionpapers/2020-1

Received December 26, 2019 **Accepted as Economics Discussion Paper** January 6, 2020
Published January 7, 2020

1 Motivation

During the past thirty years, numerous meta-analyses have been published in economics. Typically, these analyses pull a sample of empirical studies from the literature on a particular subject and then summarize or synthesize the distribution of estimates using precision weighted-means and meta-regressions.[1] Covariates or moderators in regressions include variables that describe primary data samples (e.g., country source) and methods employed by primary investigators (e.g., OLS, panel data). Categorical variables are frequently used as moderators, thereby creating subgroups for analysis. Despite the high degree of heterogeneity that exists in economic data and studies, an overwhelming majority of meta-analyses in economics employ a conditional or fixed-effect size (FES) model to summarize and analyze study results. The FES model assumes that sampled estimates represent the only population of interest, which means analytical results should not be generalized beyond these data (Borenstein 2019). In the limiting case of a common-effect model, there is only one true population effect. The usual objective of a meta-analysis is to report a weighted-mean effect size or, more commonly in economics, analyze variation in primary estimates using a meta-regression. Sources of variation in a FES model are sampling error; variation in data and methods used by primary investigators; and possibly publication bias. This characterization of a FES model was explained in detail by Hedges and Vevea (1998); see also Borenstein et al. (2009, 2010). In contrast, a model of random-effects size (RES) assumes *a priori* the primary estimates are a sample from a larger population of potential studies (Hedges 1983). The RES or plural-effects model is appropriate if primary study estimates are representative of a universe of comparable populations and the objective of the meta-analysis is to generalize in terms of a grand mean and variation around that mean (Borenstein 2019). Sometimes the universe in question is suggested by the analyst, but it is likely that readers and policymakers also are prone to generalize. However, inferences beyond the sample of estimates should not be carried out using a FES model. The population is clearer in a FES model which limits inferences, but a RES model permits unconditional inferences to broader populations (Hedges and Vevea 1998). In a RES model, variation of true effects across populations – the between-study variance – is explicitly incorporated in the analysis, thereby controlling for this additional source of dispersion. In contrast, a FES model assumes the between-study variance is zero or nearly so, which implies perfect homogeneity of true effects. Choice of model and attendant weights is therefore an important aspect of any meta-analysis.[2]

The objective of this paper to illustrate how results from a RES model can differ from those in a FES model, reflecting unobserved or random heterogeneity in a sample of estimates. An extended example

[1] A search of *EconLit* in December 2019 using the search term "meta-analysis" in the field for "Abstract" returned 1,032 citations for 1985-2019, with about half of the citations appearing during and after 2012.

[2] The terms fixed-effect and random-effects refers to different weightings used in each case and should not be confused with similar terminology used in panel data econometrics.

is employed using empirical estimates for pass-through rates for excise taxes levied on alcohol beverages; i.e., effect size estimates measure the extent to which a one cent increase in a tax causes change in retail prices of alcohol. A pass-through of unity is a fully-shifted tax and provides a convenient benchmark for the null hypothesis. Both FES and RES models allow that mean pass-throughs can be greater or less than unity, so a chief difference is the range of possible inferences. As demonstrated below there is substantial dispersion of estimates, which suggests use of a RES model. Data on pass-throughs rates are contained in 30 primary studies, which cover alcohol excise taxes and prices for beer, wine or spirits. There are 16 studies for the U.S. and 14 studies that use data for European nations, groups of countries (EU, OECD), and various other nations (Japan, South Africa). U.S. studies use data that cover alcohol taxes and prices for individual brands, cities, states or the nation. As is typical in economics, the diversity of data and econometric methods employed by primary researchers is substantial and range from simple OLS estimates based on aggregate time series (Barzel 1976) to scanner prices and panel data at the city level (Chua 2000). Non-U.S. studies also use diverse data and methods. No two studies are identical. A variety of special issues also are addressed in primary studies such as border effects arising from tax increases in one jurisdiction (Bergman and Hansen 2017) or pass-through rates that vary across the price spectrum (Shang et al. 2018). These issues are not addressed here, but methods and results are described at length in a prior narrative review (Nelson and Moran 2019). Primary estimates for special issues are ignored as the focus here is on price impacts at the beverage level or for all alcohol.

The choice of FES vs. RES models is crucial as the starting point for a meta-analysis. However, relatively few analyses in economics dwell on this distinction or inform readers that a FES analysis cannot be extended or generalized beyond the collection of primary studies (e.g., Anderson and Kichkha 2017; Button 2019). For example, Nelson and Kennedy (2009) surveyed 140 meta-analyses in environmental economics, where an out-of-sample benefit transfer was often an objective. Only 14 studies report weighted-means and only 22 studies employ a RES model for meta-regressions. The present paper addresses similarities and differences between the models and presents results for both weighted-means and meta-regressions for alcohol tax pass-through rates. The remainder of the paper is organized as follows. First, following a description of the data, weighted-means are reported for fixed and random-effects for all alcohol beverages, together with various statistical indexes that describe dispersion and heterogeneity in these data. Indexes reported include the *prediction interval* for random-effects, which is a recent addition to methods used in meta-analysis (Borenstein et al. 2010; Higgins et al. 2009; IntHout et al. 2016). Second, data are divided into subgroups with the analysis repeated for: (1) U.S. vs. not-U.S. estimates; (2) beer vs. wine-spirits estimates; and (3) published studies vs. not-published studies. Methods of analysis for subgroups includes the mixed-effects model, which assumes subgroups are fixed but estimates within each group are random. Third, results from meta-regressions are presented with moderators for data and method,

with attention also to issues of publication selection bias. Results are reported for funnel plots, Egger's test for plot asymmetry, and a cumulative meta-analysis that quantifies influence of less-precise estimates. Moderators in meta-regressions test for dispersion due to observable heterogeneity. Predicted means and confidence intervals are presented. Fourth, results are summarized with a focus on current practice in economics. Nelson and Moran (2019) concluded the average pass-through rate is not significantly different from unity. The present study is an extended sensitivity analysis of that conclusion. Beer taxes also presented the strongest possible case for overshifting and this study provides additional tests of that conjecture. Overall, both FES and RES models support a pass-through of unity across beverages, but dispersion around that value is substantial. Hence, a RES model is more appropriate for these data, particularly for purposes of value transfers and similar policy applications.

2 Fixed- vs. random-effects weighted-means: Tax pass-through rates

Effect-size estimates are pass-through rates for excise taxes on alcohol beverages. Pass-through rates are important as a component of optimal alcohol tax calculations for various countries (e.g., Griffith et al. 2019); price elasticity calculations (e.g., Cook 2007: 72); and as indicators of the incidence of such taxes.[3] A rate or effect size of unity is consistent with a competitive market, operating with a perfectly elastic supply schedule. Positive pass-through rates of less than unity also may occur due to supply-side inelasticity. Rates greater than unity are generally believed to be due to imperfectly competitive markets (Stern 1987; Weyl and Fabinger 2013), but other causes are possible (e.g., menu costs of price changes, vertical market structures). Hence, the null hypothesis in the present study is an average rate that is not statistically different from unity, but alternative values can be greater or less than unity. A focus in the empirical work is the range of values indicated by confidence intervals and prediction intervals, and not just point estimates.[4] Following an extensive literature search, 30 primary studies were located that provide estimates of pass-through rates for one or more beverages, including 12 academic articles and 18 unpublished studies in the "grey literature" (9 working papers, 5 dissertations and masters theses, 4 consulting reports).[5] Studies were published beginning in 1962 (Niskanen 1962) and ending in 2018, with

[3] Griffith et al. (2019:21) note that optimal tax calculations "assume that the pass-through of the tax is complete . . . [but] relaxing this assumption would require modifying the optimal tax problem."

[4] De Long and Lang (1995:1271) is an early statement of the need to report statistical confidence intervals and not just significance tests on point estimates. They argue that most null hypotheses in economics are rejected, so the important issue for a given model is the "ranges of parameter values that are excluded by empirical estimates."

[5] Four main literature sources were searched: *JSTOR; PubMed; IDEAS/RePEc;* and *Google Scholar*. In addition, searches were conducted using *Dissertation Abstracts; Social Science Research Network (SSRN);* and *Web of Science.* Exclusions include primary studies with combined excise and VAT taxes; undergraduate papers; and one consulting report where comparable estimates could not be extracted.

a median year of 2012. The median year for data is 1999. Sixteen studies provide evidence for the U.S. and 14 studies cover a variety of other countries including western European nations (Belgium, Denmark, Finland, France, Ireland, United Kingdom); eastern European nations (Hungary, Latvia, Slovenia); groups of countries (EU, OECD); and two other nations (Japan, South Africa). The 30 primary studies are listed below in Appendix A.

A meta-analysis requires a sample of estimates and their standard errors as a measure of "quality" or precision. The sample consists of 76 pass-through rates for alcohol taxes, including 40 estimates for beer, 9 for wine; and 27 for spirits. Due to the small sample for wine, estimates for wine and spirits are combined.[6] These estimates are selected from 24 of 30 studies but constitute only a sample of reported rates. Eight primary studies report more than 40 estimates each and two of these studies report more than 90 estimates each. Meta-analysis offers little guidance about which estimates are comparable or statistically independent, and a common criticism is that analyses often combine "apples and oranges" (Borenstein et al. 2009: 379). In order to obtain comparable and meaningful estimates, the sample was selected using the following criteria: first, estimates are excluded if standard errors are missing or if a tax elasticity is reported rather than a rate. Second, some excluded estimates are unique or disparate including a study reporting negative rates; rates for narrowly-defined price categories or beverage types (e.g., budget-priced beer); and rates for narrowly-defined geographic areas (e.g., border areas).[7] The number of estimates for narrower categories are too small to permit separate quantitative analysis. Third, estimates are selected to represent independent estimates, such as state or federal taxes but not multiple rates for state taxes in the same study. Overall, a restricted sample is more likely to cover comparable data and populations, but the sample also could reduce dispersion or publication bias.[8] However, the data are consistent with a diverse sample of effect sizes.

Formally, in a meta-analysis there are n estimates of the population effect that are assumed to be independent. In a FES model, the true effect-size is given by β and the estimate reported in the *i-th* study is denoted by Y_i, with a standard error s_i and precision $1/s_i$. The model assumes observed estimates are generated by $Y_i = \beta + e_i$, where e_i is a sampling error with mean zero and variance σ^2. Fixed-effect weights

[6] Estimating separate pass-through rates by beverage is a feature of all primary studies, although some studies consider only one or two beverages. Several major wine-producing countries do not tax wine, which limits the estimates.

[7] The absence of "negative" estimates in an empirical literature is frequently mentioned as an indicator of publication bias, but it could mean either insignificant (null) estimates or estimates with significant negative signs. In this analysis, it is the latter but the primary study in question (Hanson and Sullivan 2016) is unique in estimating a model with both sales taxes and excise taxes for two leading beer brands (Bud Light, Miller Lite). They argue that pass-through of excise taxes is negative because consumers "overreact" to tax increases.

[8] Rhodes (2012: 27) notes that "the meta-analyst defines the population . . . basically, definition of the population is driven by the need to average estimated effect sizes across studies, so the meta-analyst avoids averaging across treatments that differ greatly from each other."

are defined as $w_i = 1/s_i^2$. As summary statistics, the FES weighted-mean and variance are given by (Hedges and Olkin 1985):

$$\bar{\beta}_F = \sum w_i Y_i / \sum w_i \,, \quad \tilde{\sigma}_F^2 = 1/\sum w_i \qquad i = 1, \ldots, n \tag{1}$$

with the 95% confidence interval given by $CI_F = \bar{\beta}_F \pm t^{.05}(\tilde{\sigma}_F)$. Alternatively, the RES model assumes each estimate is a draw from a distribution of true effects. Thus, a RES analysis is designed to facilitate unconditional inferences about (non-sampled) studies that are similar but not identical to sampled studies (Hedges and Vevea 1995). Heterogeneity among underlying populations is modeled by $Y_i = \beta_0 + u_i$, where β_0 is a grand effect and u_i is an error term with mean zero and variance τ^2. Random-effects assume that observed estimates are generated by $Y_i = \beta_0 + u_i + e_i$, where errors are assumed to be independent. The estimated variance of each primary effect is given by a composite error term, $V_i^2 = s_i^2 + T^2$, where T^2 is an estimate of variation in true effects around the grand mean. Note that there is one value of T^2 for all estimates in a sample but there are several possible methods for obtaining an estimate, with method-of-moments used here (DerSimonian-Laird method). The RES weighted-mean and variance are given by equation (2), with an estimate of the grand mean denoted by $\bar{\beta}_R$ and weights $w_i^* = 1/V_i^2$:

$$\bar{\beta}_R = \sum w_i^* Y_i / \sum w_i^* \,, \quad \tilde{\sigma}_R^2 = 1/\sum w_i^* \qquad i = 1, \ldots, n \tag{2}$$

with the 95% confidence interval given by $CI_R = \bar{\beta}_R \pm t^{.05}(\tilde{\sigma}_R)$, showing the precision of the mean. The 95% prediction interval is given by $PI_R = \bar{\beta}_R \pm t^{.05}\sqrt{T^2 + \tilde{\sigma}_R^2}$, which quantifies the dispersion of the mean estimate and the (estimated) variance of true effects (Higgins et al. 2009). The variance of true effects does not approach zero as estimates increase in number.

As a third measure, Thompson and Sharp (1999) propose a "weighted least-squares" (WLS) estimator that incorporates residual heterogeneity with weights proportional to inverse effect variances.[9] Stanley and Doucouliagos (2015) suggest the WLS-estimator is superior to standard FES and RES-estimators. First, they argue that fixed-effect confidence intervals are too small if applied to unconditional inferences since the assumption is that estimates are a sample from a homogeneous population ($\tau^2 \approx 0$). Second, they argue that random-effects are sensitive to estimates of τ^2 if the sample is small or suffers from publication bias. They propose instead a WLS-estimator with weights given by $w_i' = 1/(\phi s_i^2)$, where $\phi > 1$ is a proportionality constant. The WLS-estimator produces the same mean as a FES-estimator, but its variance is different; i.e., $\tilde{\sigma}_L^2 = \phi/(\sum w_i) = \phi\tilde{\sigma}_F^2$ (Stanley and Doucouliagos 2015: 2117). An appropriate estimate of ϕ can be obtained from the mean squared error from an OLS regression of standardized effect sizes ($t_i = Y_i/s_i$) on their precision. Alternatively, a comparable measure is the H^2 statistic for overdispersion

[9] Reed (2015: 3) notes that all meta-analysis mean estimators also are weighted least-squares estimators; i.e., weighted regressions with only a constant. The WLS-version reported here is a specific variant of a FES-estimator and should not be confused with similar terminology used in econometrics.

(Thompson and Sharp 1999: 2699). The H^2 statistic is obtained by dividing Cochrane's Q-statistic by n - 1. To obtain WLS standard errors, FES standard errors are multiplied by $\sqrt{\phi}$ = H. However, because ϕ is derived from the Q-statistic, it suffers from the same imprecision that arises in estimating τ^2 from small samples (Thompson and Sharp 1999: 2699).

Table 1 summarizes the sample together with unweighted averages for several subgroups and confidence intervals. The range of estimates is substantial, arising from several possible sources including sampling error, data type, econometric methods, heterogeneous populations, outliers, and publication bias. Unweighted averages indicate pass-through rates are greater than unity and more substantial for the beer subgroup. The histogram in Figure 1 also indicates a broad range of effect sizes, although a majority of estimates are clustered in the range 0.5 to 1.5. For the full sample (n = 76), Table 2 displays results for FES, WLS, and RES-means together with statistical measures for dispersion and heterogeneity. The FES-mean for alcohol is 1.04. The 95% confidence interval is reasonably narrow and barely exceeds unity, 1.02 to 1.07. The wider confidence interval for the WLS-mean is 0.99 to 1.10. The RES-mean gives greater weight to less precise estimates.[10] The RES-mean for alcohol is 1.16 and the confidence interval is 1.09 to 1.24, suggesting on average there is overshifting of taxes. The prediction interval is 0.69 to 1.64, consistent with both under- and overshifting of taxes. Heterogeneity statistics indicate substantial dispersion in the sample:

● **Q-statistic.** The Q-statistic is a standardized measure of dispersion or weighted sum of squared deviations (WSS); i.e., deviation of each primary effect from the FES-mean, weighted by the inverse-variance of that effect and summed over all values in the sample. Q measures total dispersion around the mean and Q – df is residual or "excess dispersion," where df = (n -1) is degrees of freedom. If there is a common effect size, then Q = df. Under a null hypothesis of a common effect, Q has a chi-square distribution with n – 1 degrees of freedom. The p-value in Table 2 indicates the null is rejected for n = 76 at the 95% confidence level. Power for the Q-statistic is discussed in Borenstein et al. (2009: 272) and Rhodes (2012:45).

● **H-statistic**. H^2 = Q/(df) is the ratio of observed WSS to expected WSS if there is a common effect. Thus, H = 1 when there is homogeneity of effect sizes, which is rejected in Table 2.

● **I-squared**. This statistic equals ((Q – df)/Q) x 100% = ((H^2 -1)/H^2) x 100%, which is a ratio of excess dispersion to total dispersion. I^2 is thus analogous to 1- R^2 for regressions. In general, an I^2 less than 25% is considered low relative dispersion; 25-75% is moderate dispersion; and greater than 75% is high dispersion (Higgins et al. 2003). The Table 2 value of 82% indicates high dispersion. According to Reed (2015), values of 75-95% are common in economics studies.

● **T-squared.** This statistic is the sample estimate of the variance of true effect sizes (τ^2), which is used to assign weights in the RES model. Larger values of T or T^2 are indicative of greater dispersion. In Table 2, the standard error for T^2 is 0.019, which indicates that 95% of values are likely to lie in the range 0.018 to 0.092, with a mean of 0.055. Conventional confidence intervals are reported based on large samples.

[10] Weights in the FES and RES models are substantially different. Raw weights in the FES model vary from 625.0 for the most precise estimate to 0.36 for the least precise, with a median weight of 14.8. Relative weights vary from 9.44% to 0.005%, with a median of 0.22%. For the RES model, raw weights vary from 17.5 to 0.36, with a median weight of 8.12. Relative weights vary from 2.57% to 0.05% , with a median of 1.19%.

Table 1: Summary statistics for alcohol tax pass-through rates

Sample & statistic	Value (sd)	95% Confidence interval
Alcohol tax sample (n = 76)		
Unwt. mean rate (sd)	1.41 (0.69)	0.06 – 2.76
Median rate	1.20	
Range	0.56 – 3.84	
Unwt. mean std error (sd)	0.34 (0.30)	
Unwt. mean t-stat. (sd)	7.62 (6.55)	
U.S. subgroup (n = 45)		
Unwt. mean rate (sd)	1.45 (0.65)	0.18 – 2.72
Median	1.41	
Range	0.56 – 3.19	
Beer subgroup (n = 40)		
Unwt. mean rate (sd)	1.57 (0.81)	-0.02 – 3.16
Median	1.34	
Range	0.56 – 3.84	
Published subgroup (n = 30)		
Unwt. mean rate (sd)	1.43 (0.58)	0.29 – 2.57
Median	1.28	
Range	0.62 – 3.00	

Notes: standard deviations (sd) in parentheses.

Figure 1: Histogram of alcohol tax pass-through rates (n = 76)

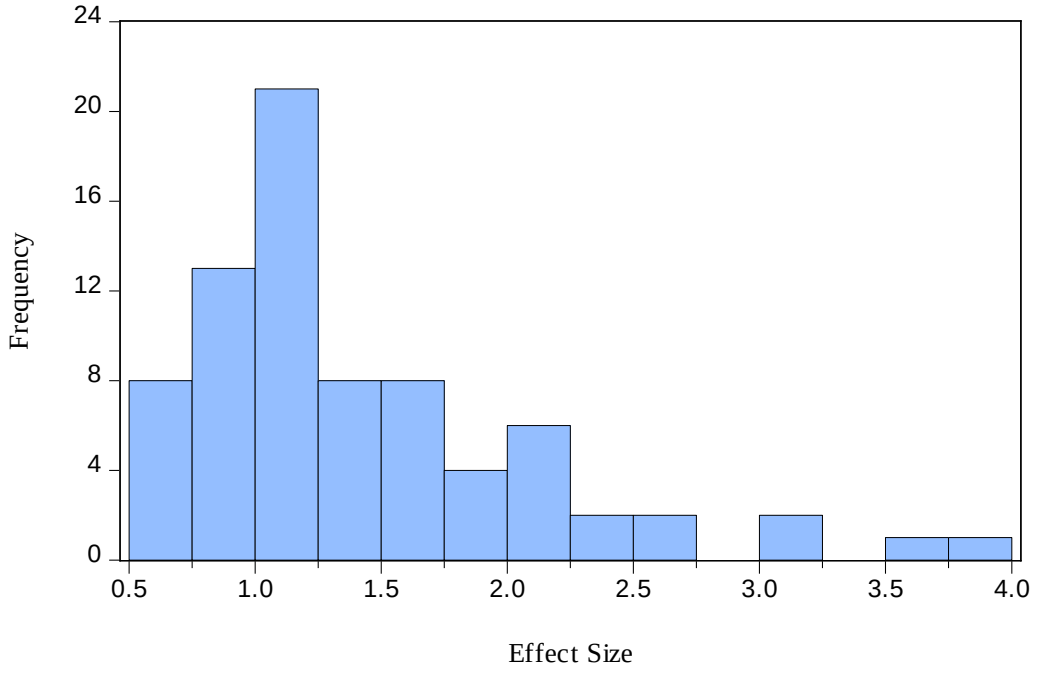


Table 2: Fixed and random-effects means – Alcohol tax pass-through rates

Sample & statistic	Value (se)	95% CI	95% PI
Alcohol taxes (n = 76)			
FES mean	1.044 (0.012)	1.020 – 1.068	
WLS mean	1.044 (0.028)	0.989 – 1.100	
RES mean	1.164 (0.038)	1.089 – 1.238	0.698 – 1.630
Q-stat. (p-value)	422.7 (0.000)		
H-stat.	2.37	2.14 – 2.62	
I-sq. %	82.3	78.3 – 85.5	
T-sq. (se)	0.055 (0.019)	0.018 - 0.092	

Notes: Method-of-moments used for random-effects T^2. All statistics computed in *Comprehensive Meta Analysis 3* (2019). Basic results checked against *Stata 16*. CI = confidence interval. PI = prediction interval. se = standard error. Formula for I-sq. confidence interval using Q is from Borenstein et al. (2009: 124). Formula for a test-based H confidence interval using Q is from Higgins and Thompson (2002: 1554). Conventional confidence interval used for T^2 based on sample sizes.

● **95% Prediction interval**. This statistic combines information about the variance of the mean effect (se^2) and the variance of true effects (T^2). For mean effects, the confidence interval (CI) quantifies the accuracy of estimated means, but the prediction interval (PI) shows the potential distribution of true effects (Borenstein et al. 2009: 131). The prediction interval is the measure of heterogeneity that best captures possible dispersion of true effects (Borenstein 2019: 100). It is a more conservative measure of results in a meta-analysis. PI range in Table 2 includes unity but does not rule out either under- or overshifting of taxes.

The RES prediction interval in Table 2 is substantial, indicating that pass-through rates for alcohol beverages can be greater or less than unity. In contrast, the confidence interval for the FES model exceeds unity by only a modest amount. The WLS confidence interval includes unity, so the three models offer different results. Although dispersion statistics are not commonly reported in meta-analyses conducted in economics, it seems likely the estimates in Table 2 is not unrepresentative of economic data generally. Most primary samples in economics reflect a wide range of data and empirical methods, making assumption of a common effect size untenable (Nelson and Kennedy 2009). If inferences are restricted conditionally as in the case of fixed-effects, the problem is partially solved. However, the danger is that analysts or readers generalize results or seek to make inferences that go beyond the conditional population such as required by policy applications. Several alternatives are available to address issues of excess dispersion and random heterogeneity. First, the analyst might choose to abandon meta-analysis and conduct instead a traditional narrative review of the literature. Second, the analyst could utilize subgroups of estimates where assumption of a common effect may be defensible. Third, the analyst can estimate meta-regressions, including covariates that control for systematic differences. This is the methodology commonly chosen in economics, although it needs to be emphasized that either FES or RES models can be applied in a meta-regression.

3 Subgroup analysis for countries, beverages, and publication status

Identifying subgroups within a sample is one possible solution to issues of excess dispersion, where the primary estimates within each subgroup share a well-defined characteristic. If estimates within a subgroup are homogeneous, use of the FES model is correct for that group. In the ideal scenario, there is homogeneity in each subgroup and inferences using means for each group are possible. Mean effects across subgroups can be compared or combined, which can add to insights from the analysis. Hence, meta-analysis of subgroups is often conducted using ANOVA methods. Further, the RES model requires calculation of T^2, but the analyst has a choice of whether to pool subgroup estimates or not. Given dispersion within subgroups, it is assumed that estimates within a subgroup are random, but subgroups are fixed; i.e., the pooled model is based on mixed-effects size (MES) that permits generalizations to comparable populations (Borenstein 2019: 196). This section performs meta-analysis for three categories of subgroups: (1) country source (U.S. vs. not-U.S. rates); (2) beverage type (beer vs. wine-spirits); and (3) publication status (published vs. not-published studies). Pooled estimates are compared to the RES-mean value in Table 2.

Table 3 displays results for subgroup analysis. First, there is more dispersion for the U.S. subgroup as indicated by RES-means. However, I^2 values for both subgroups are substantial, 74% and 88%. The prediction interval for the U.S. subgroup is larger, but Q-statistics reject homogeneity within subgroups. Second, the beer subgroup is highly dispersed as indicated by I^2 of 84%; the 95% prediction interval, 0.35 to 2.26; FES-mean compared to the RES-mean, 1.14 vs. 1.30; and T^2 equal to 0.223. The wine-spirits subgroup is somewhat less dispersed. Third, published studies also exhibit a higher degree of dispersion. The I^2 value is 87% and the 95% prediction interval is 0.57 to 1.97. RES-means for published and unpublished studies differ substantially, 1.27 compared to 1.10. Fourth, pooled MES-means do not differ much from the RES-mean in Table 2, 1.164–1.168 compared to 1.164. Fifth, many confidence intervals for subgroups in Table 2 include unity – there are 14 of 18 confidence intervals for means that include unity or are very close to that value (i.e., 1.01). Exceptions include RES intervals for U.S.-based estimates, beer, and published studies. This helps identify subgroups with larger and less precise estimates, a possible indicator of publication bias. All prediction intervals include unity, but ranges are substantial. All within-group Q-statistics reject homogeneity, which is a common outcome in similar analyses. It is well-known that tests based on Q have low power, so it is important to consider a variety of tests and procedures for random-effects. For example, $T^2 > 0$ in five of six confidence intervals. Lastly, all subgroup prediction intervals include unity, but the intervals are quite wide for U.S.-based estimates, beer, and published studies. Full-shifting of taxes is possible, but either under- or overshifting cannot be ruled out based on the evidence in Tables 2 and 3.

Table 3: Fixed and random-effects meta-analysis – Subgroup analysis

Subgroup & statistic	Value (se)	95% CI	95% PI
US estimates (n = 45)			
FES mean	1.056 (0.023)	1.011 – 1.100	
WLS mean	1.056 (0.045)	0.968 – 1.144	
RES mean	1.257 (0.059)	1.141 – 1.372	0.698 – 1.816
Within Q-stat. (p-value)	167.4 (0.000)		
H-stat.	1.95	1.69 – 2.26	
I-sq. %	73.7	64.9 – 80.3	
T-sq,	0.078 (0.051)	-0.02 – 0.178	
Not-US estimates (n = 31)			
FES mean	1.039 (0.015)	1.010 – 1.067	
WLS mean	1.039 (0.044)	0.953 – 1.125	
RES mean	1.099 (0.052)	0.996 – 1.201	0.641 – 1.557
Within Q-stat. (p-value)	254.8 (0.000)		
H-stat.	2.91	2.53 – 3.35	
I-sq. %	88.2	84.4 – 91.1	
T-sq.	0.052 (0.021)	0.011 – 0.093	
Pooled MES mean	1.168 (0.039)	1.091 – 1.245	0.702 – 1.634
Beer estimates (n = 40)			
FES mean	1.137 (0.031)	1.076 – 1.199	
WLS mean	1.137 (0.078)	0.984 – 1.290	
RES mean	1.304 (0.059)	1.188 – 1.420	0.351 – 2.257
Within Q-stat. (p-value)	245.2 (0.000)		
H-stat.	2.51	2.19 – 2.87	
I-sq. %	84.1	79.2 – 87.8	
T-sq.	0.223 (0.096)	0.035 – 0.411	
Wine-spirits (n =36)			
FES mean	1.027 (0.013)	1.001 – 1.053	
WLS mean	1.027 (0.028)	0.972 – 1.082	
RES mean	1.063 (0.050)	0.965 – 1.161	0.738 – 1.388
Within Q-stat. (p-value)	167.0 (0.000)		
H-stat.	2.18	1.87 – 2.54	
I-sq. %	79.0	71.6 – 84.6	
T-sq.	0.025 (0.011)	0.003 – 0.047	
Pooled MES mean	1.164 (0.038)	1.089 – 1.239	0.698 – 1.630
Published studies (n = 30)			
FES mean	1.051 (0.024)	1.004 – 1.097	
WLS mean	1.051 (0.067)	0.920 – 1.182	
RES mean	1.273 (0.062)	1.152 – 1.394	0.572 – 1.974
Within Q-stat. (p-value)	224.6 (0.000)		
H-stat.	2.78	2.40 – 3.22	
I-sq. %	87.1	82.7 – 90.4	
T-sq.	0.124 (0.062)	0.002 – 0.246	

Table 3: (continued)	**Value (se)**	**95% CI**	**95% PI**
Unpublished (n = 46)			
FES mean	1.041 (0.014)	1.013 – 1.069	
WLS mean	1.041 (0.029)	0.984 – 1.098	
RES mean	1.097 (0.050)	0.999 – 1.195	0.722 – 1.471
Q-stat. (p-value)	198.0 (0.000)		
H-stat.	2.10	1.83 – 2.41	
I-sq. %	77.3	70.0 – 82.8	
T-sq.	0.034 (0.016)	0.003 – 0.065	
Pooled MES mean	1.167 (0.039)	1.091 – 1.243	0.701 – 1.633

Notes: See notes for Table 2. RES-means calculated using a mixed-effects size (MES) model for subgroups. The estimate of the sample variance (T^2) is assumed to be the same for all subgroups; i.e., a value is computed within subgroups and then pooled across subgroups using fixed-effects to obtain the estimate. RES-means for beer and wine-spirits differ from those in Table 5 due to assumptions of the MES model. Between-group Q-statistics (p-value) for fixed and random-effects are: (1) country source, 4.10 (0.043) and 11.2 (0.001); (2) beverage type, 10.5 (0.001) and 9.66 (0.002); and (3) publication status, 0.119 (0.730) and 4.90 (0.027).

4 Detecting and reducing publication bias: Tax pass-through rates

Publication or reporting bias is now recognized as a major issue in many areas of scientific inquiry, including economics (Christensen and Miguel 2018; Costa-Font et al. 2013; De Long and Lang 1992; Card and Krueger 1995; Ioannidis et al. 2017). Bias is important because it can lead to incorrect conclusions for a narrative review or quantitative analysis, including false positives or Type 1 errors. Publication bias is generally understood to have two important dimensions, both of which tend to exaggerate reported estimates of effect sizes. First, non-publication of weak, null or contrary results, which is known as the "file-drawer problem." Second, a tendency for scientific journals to publish statistically significant results and avoid "negative results," resulting in data-mining (aka "p-hacking"); i.e., significant findings are more likely to be published causing bias as researchers mine their data to find publishable estimates. The first problem can be addressed in part by a thorough search of the literature, which is greatly aided by web-based resources. The second issue requires, first, detecting publication bias in a sample of primary estimates and, second, application of methods to reduce bias and obtain corrected estimates for mean effect sizes. This section presents two methods to detect publication bias (funnel plots, Egger's funnel asymmetry test) and three methods to reduce publication bias (trim-and-fill, cumulative meta-analysis, meta-regressions). Results are reported for the full sample and selectively by beverage for fixed and random-effects models. Meta-regression results include predicted means and confidence intervals.

Figure 2 shows funnel plots of effect sizes plotted against standard errors (2a) or precision (2b), with a vertical line for the FES-mean. In the absence of publication bias, effect sizes will be distributed symmetrically about the mean. Funnel-shaped lines represent 95% confidence intervals for between-study variation; i.e., in the absence of heterogeneity 95% of all estimates would lie within these lines. Both plots

Figure 2a: Funnel plot of pass-through rates and standard errors (n = 76)

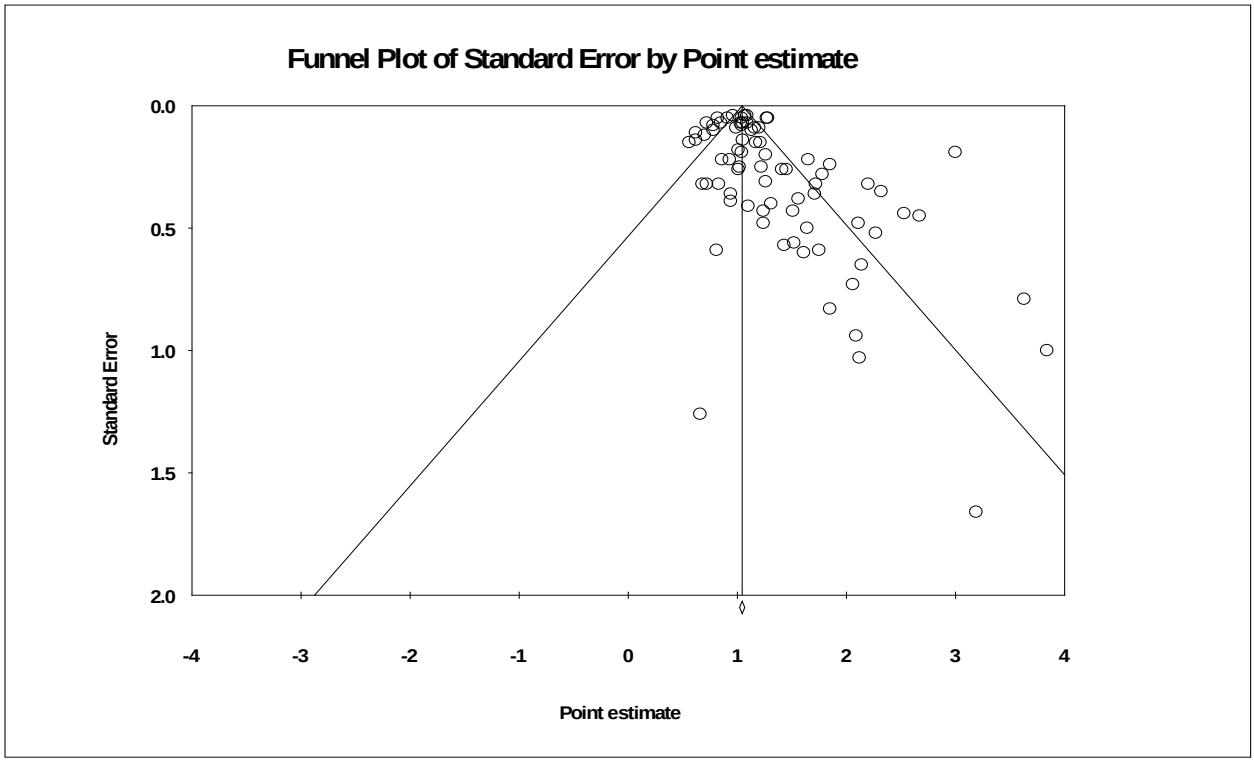


Figure 2b: Funnel plot of pass-through rates and precisions (n = 76)

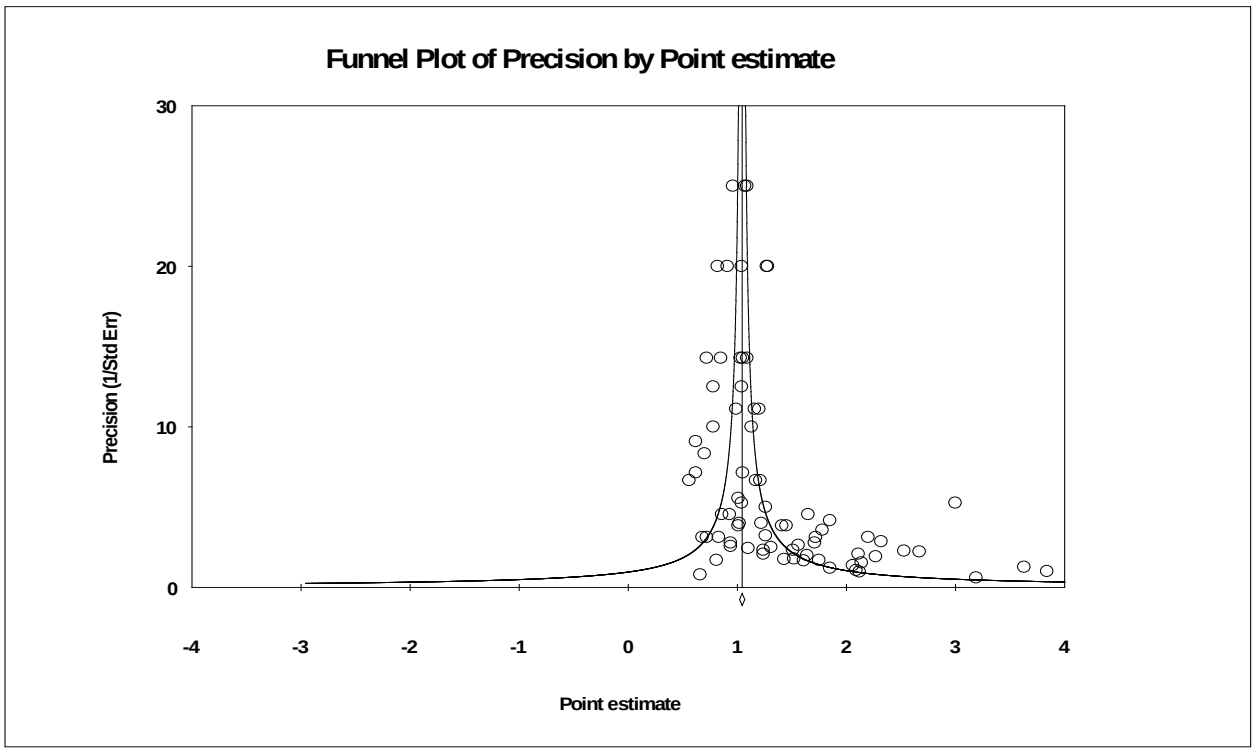


Table 4: Funnel asymmetry regression tests for publication bias (n = 76)

Funnel asymmetry test (FAT-PET)	**Value (se)**	**95% CI**
Intercept	1.241 (0.307)	0.639 – 1.843
FES precision is 1/se	0.948 (0.045)	0.860 – 1.036
R-sq.	0.886	
F-stat. (p-value)	575.8 (0.000)	
Precision effect test (PEESE)		
Std. error variable	1.871 (0.419)	1.050 – 2.692
FES precision is 1/se	1.022 (0.035)	0.953 – 1.091
R-sq.	0.951	
F-stat. (p-value)	508.5 (0.000)	
Lin-Chu test		
Intercept	2.087 (0.333)	1.434 – 2.740
RES precision is 1/(se + T)	0.513 (0.108)	0.301 – 0.725
R-sq.	0.164	
F-stat. (p-value)	14.47 (0.000)	

Notes: Dependent variable is standardized effect size. Heteroskedastic-consistent standard errors in parentheses.

suggest "missing" values to the left of the mean, but this could be due to publication bias, study-level heterogeneity or causal factors such as imperfect markets for alcohol. The plots also indicate that larger estimates tend to have larger standard errors reflecting natural heteroskedasticity. More precise estimates in Figure 2 are clustered at the top of funnel plots in the range 0.5 to 1.5.

An alternative to visual inspection of plots is Egger's regression test for asymmetry, obtained by regressing standardized effect sizes ($t_i = Y_i/s_i$) on corresponding FES precision ($1/s_i$) and an intercept. In economics, this is referred to as a *funnel asymmetry test* (FAT); see Stanley and Doucouliagos (2012: 61). In the presence of publication bias, the intercept is significantly different from zero and the slope coefficient is a biased-adjusted estimate of the FES-mean; i.e., a *precision-effect test* (PET). Further, Stanley and Doucouliagos (2012: 65) also propose adding effect-size standard errors (s_i) to the FAT regression in order to capture non-linearities, which they label the *precision effect estimate with standard error* (PEESE). It should be emphasized that standard errors are themselves estimates, so all meta-regression tests are possibly biased (Stanley 2008). Lin and Chu (2018) propose another modified version for samples with significant heterogeneity, where the effect size is divided by the RES standard error and the covariate is the RES-adjusted precision. Table 4 (above) displays results for the three tests for asymmetry, where indicators of bias are intercept terms and the slope coefficient on the standard error variable in the PEESE regression. All three terms are significantly positive, consistent with selection bias. Precision coefficient values and confidence intervals for PET and PEESE regressions provide support for the null hypothesis of unity. Lin-Chu's test performs poorly, although the dependent variable is different.

Having detected the possible presence of publication bias, the next issue is what to do about it or how to correct mean effects for bias in FES and RES models. Three main methods are available. The first is the "trim-and-fill" procedure due to Duval and Tweedie (2000), which is an iterative procedure that identifies those estimates responsible for asymmetry in Figure 2 and fills the plot with mirror-images of the outlying estimates. New mean values are computed using a filled-sample of estimates. Both FES and RES-adjusted means can be recomputed using either unadjusted mean as a base to detect missing values, but FES base means are traditionally used and employed here as well. The following results are obtained for trim-and-fill:

- **Trim-and-fill FES-means.** The adjusted FES-mean for all alcohol beverages is 1.018 (0.026), with a 95% confidence interval of 0.967–1.069. Adjusted FES-mean for beer is 1.012 (0.030), with a 95% confidence interval of 0.953–1.071. Adjusted FES-mean for wine-spirits is 1.022 (0.013), with a 95% confidence interval of 0.996–1.048. All three confidence intervals include unity, consistent with the null hypothesis. Mean values are close to one.

- **Trim-and-fill RES-means.** The adjusted RES-mean for all alcohol beverages is 1.019 (0.042), with a 95% confidence interval of 0.937 to 1.101. Adjusted RES-mean for beer is 1.063 (0.104), with a 95% confidence interval of 0.859–1.267. Adjusted RES-mean for wine-spirits is 1.008 (0.038), with a 95% confidence interval of 0.934–1.082. The three confidence intervals include unity, consistent with the null hypothesis. Mean values are close to one, except beer.

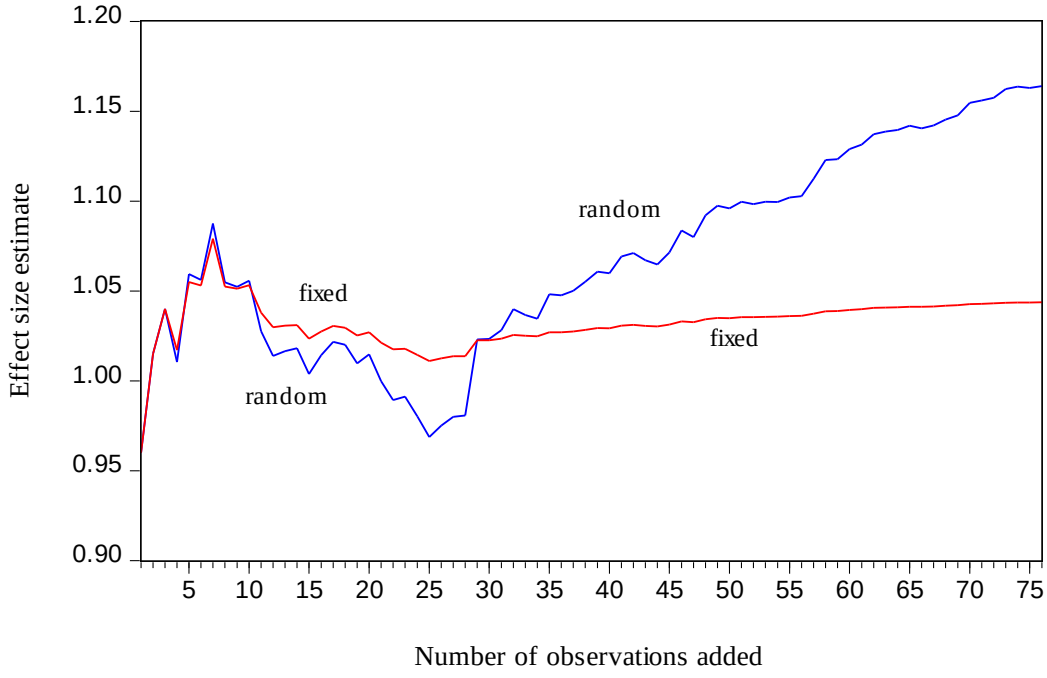
The second method for reducing publication bias is based on a cumulative meta-analysis (Borenstein et al. 2009: 287). Effect estimates are ordered in ascending size of standard errors. Sequentially, estimates are included in the sample yielding cumulative means that vary as less precise values are added. In the presence of selection bias, the expectation is that mean values will increase as the sample of estimates is expanded to include less precise values. Table 5 and Figure 3 show the results for three cumulative meta-analyses for all alcohol and each beverage. Four or five cumulative means and relative weights are reported for each array. For all alcohol beverages, the FES-mean rises only slightly but most of the weight in the analysis is on a small portion of the total sample. The first 15 estimates have a relative weight of 79%. Results for the RES model are revealing as the last 21 estimates increase the mean from 1.07 to 1.16. The first three RES confidence intervals include unity, but the final interval (n = 76) does not. The first 30 estimates yield equal mean values for FES and RES models, 1.023, and then diverge as shown in Figure 3. For beer, FES and RES-means for the first 20 estimates are 1.085 and 1.216. The sample of beer estimates is skewed toward larger and less precise values, which has an impact on the RES-mean. In contrast, mean values for wine-spirits for the first 20 estimates are 1.022 and 1.005, indicating pass-through rates that barely exceed unity regardless of model. Overall, a cumulative analysis suggests average rates of unity, except possibly beer. For all alcohol, the 21 least precise estimates strongly affect the RES-mean. For beer,

Table 5: Cumulative meta-analysis results

Beverage & model	Cumulative mean (se)	Cumulative wt. %	95% CI
All – FES model			
First 15 estimates	1.023 (0.014)	78.7	0.996 – 1.051
First 30	1.023 (0.013)	94.5	0.998 – 1.047
Median – First 38	1.028 (0.013)	96.8	1.004 – 1.053
First 45	1.031 (0.012)	98.0	1.007 – 1.056
All 76	1.044 (0.012)	100.0	1.020 – 1.068
All – RES model			
First 15 estimates	1.004 (0.040)	37.2	0.925 – 1.083
First 30	1.023 (0.042)	67.3	0.942 – 1.105
Median – First 38	1.055 (0.039)	78.0	0.978 – 1.132
First 45	1.071 (0.038)	85.2	0.997 – 1.146
All 76	1.164 (0.038)	100.0	1.089 – 1.239
Beer – FES model			
First 10 estimates	0.959 (0.036)	75.6	0.888 – 1.030
Median – First 20	1.085 (0.033)	91.4	1.020 – 1.149
First 30	1.123 (0.032)	98.4	1.061 – 1.186
All 40	1.137 (0.031)	100.0	1.076 – 1.199
Beer – RES model			
First 10 estimates	0.935 (0.074)	36.5	0.790 – 1.081
Median – First 20	1.216 (0.107)	66.8	1.006 – 1.425
First 30	1.320 (0.096)	90.2	1.131 – 1.509
All 40	1.395 (0.094)	100.0	1.211 – 1.579
Wine-spirit FES model			
First 10 estimates	1.038 (0.015)	76.4	1.008 – 1.068
Median -- First 18	1.021 (0.014)	96.9	0.995 – 1.048
First 30	1.024 (0.013)	99.7	0.998 – 1.050
All 36	1.027 (0.013)	100.0	1.001 – 1.053
Wine-spirit RES model			
First 10 estimates	1.027 (0.051)	49.1	0.927 – 1.127
Median -- First 18	0.998 (0.039)	82.6	0.922 – 1.075
First 30	1.019 (0.037)	97.6	0.947 – 1.091
All 36	1.040 (0.037)	100.0	0.967 – 1.112

Notes: Effect size estimates arrayed in ascending order from smallest to largest standard errors and added sequentially to the sample.

Figure 3: Trace of cumulative fixed and random-effects mean (max n = 76)



the RES-mean rises sharply after the first 10 estimates. Similar patterns are not found for FES-means or the wine-spirits subgroup. Except beer, median weighted-rates are approximately unity: 1.028 and 1055; and 1.021 and 0.998. The beer rates are 1.085 and 1.216. A cumulative analysis also illustrates potential problems inherent in both fixed and random-effects models. The FES model gives substantial weight to more precise estimates, which may not be representative of the population of interest or could be due to outliers. The RES model gives more weight to less precise estimates, but weak precision can be due to publication selection bias. Estimation of the between-study variance also may limit a RES model.

The third method for reducing bias is a meta-regression analysis (MRA) that seeks to explain systematic heterogeneity and simultaneously correct or reduce publication bias. However, there is division between economists and non-economists on the appropriate regression model. In the first instance, Stanley and Doucouliagos (2012) recommend using a fixed-effect model estimated by weighted least-squares with weights based on inverse variances. They argue " . . . because the standard error, or precision, is always one of the independent variables in our MRA models, a random-effects model is likely to be invalid" (Stanley and Doucouliagos 2012: 64). This argument is valid, first, if covariates can explain a substantial portion of observed variation in effect sizes.[11] However, as shown below, this is not the case for alcohol tax pass-

[11] Konstantopoulos and Hedges (2019: 246) argue "If a fixed-effects model explains all of the variation in effect-size parameters, the (fixed-effects) model is appropriate . . . if differences between studies that lead to differences in effects are *not* regarded as random . .. then fixed-effects methods are appropriate" (emphasis in the original).

through rates. Second, it must be the case that analysis is restricted to the current set of estimates and there is no desire to make unconditional inferences about broader populations. This is the standard argument in favor of random-effects. Third, it is not explained why FES weights are independent of moderators; e.g., variance inflation factors (VIF) reported in tables below are slightly smaller for RES regressions. Fourth, Stanley and Doucouliagos (2012: 85) also argue that a regression analysis should incorporate the variance as an explanatory variable (as well as the standard error). However, both variables are subject to bias (Vevea et al. 2019: 395). Fifth, as shown in the cumulative analysis, a weakness of the FES model is that considerable weight might be given to a small portion of the sample, further limiting inferences from a regression analysis. As a result of these several considerations, other commentators recommend a random-effects model (e.g., Borenstein et al. 2009: 196; Borenstein 2019: 229; Thompson and Higgins 2002), with Knapp-Hartung adjustment of standard errors on RES regression coefficients (Borenstein 2019: 36).[12] Both FES and RES regressions are reported for all alcohol and selectively by beverage. In the case of fixed-effects, a Q-test is used to determine if regression estimates are consistent with a FES regression model.

Meta-regressions can be used to control for both differences in basic data and differences in methods employed by researchers. In the case of data by beverage or country, categorical variables allow generalization to specific populations. Controls for method are observational and thus consistent with assumptions of the FES model (Rhodes 2012). Predicted means from regressions are reported to test the null hypothesis of a unitary pass-through. Weighted meta-regressions are obtained using *Comprehensive Meta Analysis v3* software (Biostat 2015; Borenstein et al. 2015), supplemented with *Stata 16* for predicted means. Regression results are reported for seven covariates with the following expectations regarding coefficient signs (binary dummy variables, except the standard error variable):

- **Standard error variable.** Expected to be significantly positive in the presence of publication bias; i.e., larger standard errors are associated with larger effect sizes.

- **Wine-spirits variable.** Subgroup analysis indicated that mean estimates for beer might be significantly larger than those for wine and spirits. Expected coefficient sign is negative. A **wine-only dummy** is defined for the nine primary observations for wine.

- **Not-US variable.** Subgroup analysis indicated that not-US estimates might differ systematically from US estimates, with expectation of a negative sign.

- **Published study variable.** Publication bias might appear as a tendency for published rate estimates to exceed those found in grey literatures, with a positive expected sign.[13]

[12] Knapp-Hartung standard errors are larger than conventional standard errors provided by, say, *Stata 16*, so meta-regressions reported for the RES model are conservative in this dimension; see Konstantopoulos and Hedges (2019).

[13] Paldam (2015) argues that top journals in economics contain more innovations and therefore are likely to report empirical results that are more variable but not necessarily larger. Lower-ranked journals contain more replication.

● **OLS-estimate variable.** Simple OLS generates effect-size estimates under classical assumptions that might be incorrect. Direction of bias is difficult to ascertain for all possible violations of classical assumptions. The expected sign is uncertain.

● **Micro-data variable.** Because estimates based on micro-data use larger sample sizes, expectations are that these estimates have smaller standard errors and potentially smaller effect sizes, *ceteris paribus*. The expected coefficient sign is negative.

● **Real data variable.** Some primary estimates are based on inflation-adjusted prices and taxes, but others are not. Nelson and Moran (2019) report that primary studies estimating both real and nominal models do not find substantially different results. The expected outcome is an insignificant coefficient with an uncertain sign.

Table 6 displays random and fixed-effects meta-regressions for all alcohol. All six coefficients for the standard error variable are significantly positive. The 95% CIs for bias-corrected intercepts include unity. Bias-corrected values for predicted means are about 0.90 and confidence intervals (last row) include unity or are close to that value in regression (5). Subgroup dummies for country source, beverage type, and publication status are never significant. Moderators for data and method are not significant, except the OLS dummy in regression (6). However, Q-tests for heterogeneity are significant, which indicates that estimates are not consistent with assumptions of the FES model; i.e., the true effect size probably varies across estimates. While predicted mean rates are close to unity, regressions (1) and (4) suggest other corrections for publication bias also can provide accurate values for adjusted rates (e.g., Egger-FAT regressions, trim-and-fill). Meta-regression results, however, provide a stronger test. Table 7 displays regression results by beverage. Publication status results for beer are consistent with positive bias, but a positive sign for micro-data is inconsistent with expectations. The predicted mean for beer in regression (2) is 0.89 and the confidence interval includes unity. The FES regression for beer possibly overcorrects for selection bias. Moderator results for wine-spirits are all insignificant, except the standard error variable. The predicted mean in regression (4) is 0.91 and the confidence interval includes unity. Q-statistics in Table 7 reject the FES model and favor the RES model for beer and wine-spirits.

Overall, the regression results in Tables 6 and 7 are consistent with an approximate pass-through rate of unity regardless of beverage. Confidence intervals, however, consistently allow for under- and overshifting of taxes, so caution is dictated with use of point values only. For these data, meta-regression results also provide support for the RES model as tests based on Q-statistics reject the FES model. The possible overshifting of beer **t**axes is identified as a topic for further research, although predicted means in Table 7 are consistent with full pass-through of beer taxes. These results are obtained using a broader set

Table 6: Meta-regressions – Alcohol tax pass-through rates

Variable	**(1) Alcohol**	**(2) Alcohol**	**(3) Alcohol**	**(4) Alcohol**	**(5) Alcohol**	**(6) Alcohol**
Intercept	0.948 (0.105)*	0.859 (0.131)*	0.866 (0.162)*	0.934 (0.044)*	0.922 (0.049)*	1.003 (0.052)*
Std. error variable	1.484 (0.322)*	1.538 (0.348)*	1.445 (0.341)*	1.271 (0.188)*	1.345 (0.193)*	1.289 (0.196)*
Wine-spirits = 1	-0.078 (0.101)	-0.089 (0.102)	-0.038 (0.112)	0.013 (0.039)	0.003 (0.039)	0.013 (0.040)
Not-US =1	--	0.067 (0.108)	--	--	0.038 (0.029)	--
Published = 1	--	0.123 (0.098)	--	--	-0.044 (0.030)	--
OLS = 1	--	--	-0.062 (0.144)	--	--	-0.208 (0.040)*
Micro-data = 1	--	--	0.100 (0.127)	--	--	-0.043 (0.033)
Real data = 1	--	--	0.068 (0.123)	--	--	0.019 (0.042)
Q-stat. (p-value)	--	--	--	366.4 (0.000)	363.6 (0.000)	331.0 (0.000)
T-sq.	0.048	0.054	0.050	--	--	--
R-sq.	0.274	0.300	0.302	0.133	0.140	0.233
Smpl. size n	76	76	76	76	76	76
Model	random	random	random	fixed	fixed	Fixed
95% CI for intercept	0.74 – 1.16	0.60 – 1.12	0.54 – 1.19	0.85 – 1.02	0.83 – 1.02	0.90 – 1.10
Pred. mean rate	0.912 (0.060)*	0.893 (0.070)*	0.920 (0.065)*	0.941 (0.028)*	0.922 (0.031)*	0.934 (0.032)*
95% CI for pred. mean	0.80 – 1.03	0.76 – 1.03	0.79 – 1.05	0.89 – 1.00	0.86 – 0.98	0.87 – 1.00

Notes: Knapp-Hartung standard errors in parentheses for RES coefficients. Asterisks indicate significance at the 95% confidence level compared to zero. Random-effects estimated using method-of-moments for the between-study variance. R-sq. values based on weighted least-squares. Predicted means set the value of the standard error variable equal to zero and all other covariates at their sample means. Variance inflation factors (VIF) for standard error variable in random-effects regressions are 1.13, 1.35, and 1.28. For fixed-effects regressions, VIFs are 1.29, 1.36, and 1.39.

Table 7: Meta-regressions by beverage – Alcohol tax pass-through rates

Variable	(1) Beer	(2) Beer	(3) Beer	(4) Wine-spirit	(5) Wine-spirit	(6) Wine-spirit
Intercept	0.959 (0.162)*	0.400 (0.209)	0.449 (0.081)*	0.912 (0.055)*	0.936 (0.083)*	0.981 (0.030)*
Std. error variable	1.498 (0.468)*	1.694 (0.424)*	2.234 (0.273)*	1.185 (0.368)*	1.253 (0.369)*	0.882 (0.282)*
Wine-only = 1	--	--	--	-0.006 (0.121)	0.063 (0.132)	0.035 (0.065)
Published = 1	--	0.558 (0.172)*	0.398 (0.065)*	--	-0.122 (0.096)	-0.110 (0.039)*
Micro-data = 1	--	0.427 (0.176)*	0.203 (0.065)*	--	0.005 (0.084)	0.022 (0.029)
Q-stat. (p-value)	--	--	151.9 (0.000)	--	--	148.8 (0.000)
T-sq.	0.177	0.160	--	0.026	0.027	--
R-sq.	0.212	0.451	0.380	0.257	0.298	0.109
Smpl. size n	40	40	40	36	36	36
Model	random	random	fixed	random	random	Fixed
95% CI for intercept	0.64 – 1.28	-0.01 – 0.81	0.29 – 0.61	0.80 – 1.02	0.77 – 1.10	0.92 – 1.04
Pred. mean rate	0.959 (0.162)*	0.873 (0.145)*	0.720 (0.057)*	0.911 (0.057)*	0.901 (0.059)*	0.952 (0.027)*
95% CI for pred. mean	0.64 – 1.28	0.59 – 1.16	0.61 – 0.83	0.80 – 1.02	0.79 – 1.02	0.90 – 1.00

Notes: Knapp-Hartung standard errors in parentheses for RES coefficients. Asterisks indicate significance at the 95% confidence level compared to zero. Random-effects models estimated using method-of-moments for the between-study variance. R-sq. values based on weighted least-squares. Predicted means set the value of the standard error variable equal to zero and all other covariates at their sample means. Variance inflation factors (VIF) for standard error variable in beer regressions are 1.00, 1.02, and 1.06. For wine-spirits regressions, VIFs are 1.10, 1.11, and 1.12.

of methods and tests than previously reported. It is reassuring that RES methods – greater weight for less precise studies – produce much the same results as FES methods, but apply to a broader set of possible circumstances. Summing-up the results in this section, meta-analysis using both models and a variety of tests establish the following: (1) funnel plots and asymmetry tests indicate that publication bias probably affects the sample of estimates; (2) trim-and-fill estimates for bias-adjusted means are close to unity, regardless of beverage type or model; (3) cumulative meta-analysis yields adjusted means close to unity, except for beer; (4) meta-regressions yield a RES predicted mean for alcohol of 0.89-0.91; beer, 0.87-0.96; and wine-spirits, 0.90-0.91; and (5) for alcohol and both beverages, RES predicted mean confidence intervals include unity. The null hypothesis is not rejected regardless of beverage. Finally, tests based on Q-statistics reject the FES model and R-square statistics are 0.30 or smaller in FES regressions.

5 Discussion

Fixed-effect and random-effects models represent different approaches to analyzing and synthesizing data with meta-analysis. Although the models use similar techniques, random-effects is more general and represents a more conservative approach to research synthesis. Both models use weighted-means to estimate a population effect, but the RES model allows the true mean to vary from study to study with the global mean being one of the parameters of interest. Both models allow for heterogeneity between studies, but the RES model assumes some portion is not easily identifiable in the form of moderating covariates. Both models allow for confidence intervals about mean values, but the RES model also allows calculation of prediction intervals that present heterogeneity of possible true effects on the same scale as observed effects. In the present study, an average pass-through rate of unity is not rejected based on a variety of tests, but prediction intervals in Tables 2 and 3 are about 0.7–1.6 on average. Both under- and overshifting of alcohol taxes is possible given existing data. However, predicted means and confidence intervals from meta-regressions in Tables 6 and 7 are generally consistent with full pass-through of alcohol taxes. In contrast, several widely-cited primary studies conclude there is overshifting of taxes (Kenkel 2005; Young and Bielinska-Kwapisz 2002). Future research should emphasize confidence intervals as a test of pass-through rates, with the null hypothesis being a rate of unity rather than zero.

Where there is a desire by analysts or users of a meta-analysis to generalize or make inferences about similar populations, choice of model is important. It is generally held that this should be done initially, rather than as part of a sensitivity analysis (Borenstein 2019; Hedges and Vevea 1998).[14] However, Anderson and Kichkha (2017) argue that meta-analyses in economics are subject to subjectivity on the part of analysts, including selection of studies, selection of covariates, choice of weights, and interpretation of results.[15] The present study has emphasized choice of weights but in the context of a wider set of procedures and tests than commonly used in economics. Anderson and Kichkha (2017) also characterize meta-regression analysis as akin to model specification searches (i.e., data mining), so they claim results of significance have little importance and are potentially misleading as a guide for primary researchers. However, confidence and prediction intervals reported in this paper provide necessary support to

[14] For example, estimates for value of a statistical life (VSL) is the most widely-studied area in economics using meta-analysis, with at least 15 meta-analyses (Doucouliagos et al. 2012; Viscusi 2018). The objective in both primary studies and meta-analyses might be a value-transfer for use in a benefit-cost calculation, so it is important to consider the robustness of FES models for such inferences. For additional discussion, see Nelson (2015) and more generally Bergstrom and Taylor (2006); Johnston et al. (2015); and Lindhjem and Navrud (2008).

[15] In any meta-analysis there are numerous substantive judgements to be made including several illustrated here, such as selection of primary studies, sample of estimates, selection of effect size, and choice of weights. These procedures are a common source of criticism and not unique to Anderson and Kichkha (2017). The "practice of meta-analysis" is discussed more thoroughly in other sources, especially the non-economics literature; see, for example, Borenstein et al. (2009: 377); Card (2012); Cooper et al. (2019); Nelson and Kennedy (2009); Ringquist (2013); and Rhodes (2012).

significance levels for point estimates. Some of the criticisms raised by the authors can be addressed by conducting a narrative review in conjunction with a meta-analysis. As a reaction to these and other criticisms of meta-analysis, the present study provides a summary of methods that go beyond exploratory FES meta-regressions. The study finds that economic data for alcohol tax pass-through rates is highly dispersed and heterogenous, which supports application of the RES model. For economic data and econometric estimates, this does not seem an unusual finding at all.

Reed (2015) conducts Monte Carlo simulations to investigate performance of five different meta-analysis estimators, including FES, WLS, RES, FAT-PET, and PEESE estimators. Publication bias is introduced to the simulations by dropping some studies from the distribution of true effects, which can be due to either statistical insignificance or wrong-signed estimates. The true effect can be zero or some positive number. Reed (2015: 19) finds a RES-estimator is often the most biased of the five estimators, due to weight given to less precise estimates, although it can be more efficient. This is a cost to be paid for the additional source of uncertainty introduced by random effects. However, Reed also points out that simulation exercises cannot fully capture the complexity of data commonly available to meta-analysts, including primary studies and estimates carried out using different econometric methods for different countries, data sets, and time periods. Reed (2015: 37) recommends that meta-analysts should report results for different estimators, including weighted-means that do not correct for publication bias. The present study illustrates this approach using a concrete example. It is reassuring that different estimators and procedures tend to arrive at the same conclusion regarding true effect sizes, but this required an extended sensitivity analysis that includes weighted-means and meta-regressions as well as a variety of tests using dispersion and heterogeneity statistics.

Acknowledgements No financial funding or other forms of support were received in connection with the conception or preparation of the paper. There are no competing interests to declare.

References

Anderson, R.G., and A. Kichkha. 2017. Replication, meta-analysis, and research synthesis in economics. *American Economic Review: Papers and Proceedings* 107 (5): 56-59.

Barzel, Y. 1976. An alternative approach to the analysis of taxation. *Journal of Political Economy* 84 (6): 1177-1197.

Bergman, U.M., and N.L. Hansen. 2017. Are excise taxes on beverages fully passed through to prices? The Danish experience. Discussion paper. Copenhagen, DK: University of Copenhagen. http://web.econ.ku.dk/okombe/BergmanHansen.pdf

Bergstrom, J.C., and L.O. Taylor. 2006. Using meta-analysis for benefit transfer: Theory and practice. *Ecological Economics* 60 (2): 351-360.

Biostat, Inc. 2015. *Comprehensive Meta Analysis Version 3.0*. Englewood Cliffs, NJ: Biostat, Inc.

Borenstein, M. 2019. *Common Mistakes in Meta-Analysis and How to Avoid Them*. Englewood, NJ: Biostat, Inc.

Borenstein, M., L.V. Hedges, J.P.T. Higgins, and H.R. Rothstein. 2009. *Introduction to Meta-Analysis*. Chichester, UK: Wiley.

Borenstein, M., L.V. Hedges, J.P.T. Higgins, and H.R. Rothstein. 2010. A basic introduction to fixed-effect and random-effects models of meta-analysis. *Research Synthesis Methods* 1 (2): 97-111.

Borenstein, M., L. Hedges, J. Higgins, and H. Rothstein. 2015. Regression in meta-analysis. Working paper. Englewood, NJ: Biostat. https://www.meta-analysis.com/

Button, K. 2019. The value and challenges of using meta-analysis in transportation economics. *Transport Reviews* 39 (3): 293-308.

Card, D., and A.B. Krueger. 1995. Time-series minimum-wage studies: A meta-analysis. *American Economic Review* 85 (2): 238-243.

Card, N.A. 2012. *Applied Meta-Analysis for Social Science Research*. New York: Guilford Press.

Christensen, G., and E. Miguel. 2018. Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature* 56 (3): 920-980.

Chua, J.D. 2000. *Essays in Public Economics*. Ph.D. dissertation. Cambridge, MA: Harvard University. http://www.worldcat.org/title/essays-in-public-economics/oclc/77070465

Cook, P.J. 2007. *Paying the Tab: The Costs and Benefits of Alcohol Control*. Princeton, NJ: Princeton University Press.

Cooper, H., L.V. Hedges, and J.C. Valentine (eds). 2019. T*he Handbook of Research Synthesis and Meta-Analysis*, 3rd edn. New York: Russell Sage Foundation.

Costa-Font, J., A. McGuire, and T. Stanley. 2013. Publication selection in health policy research: The winner's curse hypothesis. *Health Policy* 109 (1): 78-87.

De Long, J.B., and K. Lang. 1992. Are all economic hypotheses false? *Journal of Political Economy* 100 (6): 1257-1272.

Doucouliagos, C., T.D. Stanley, and M. Giles. 2012. Are estimates of the value of a statistical life exaggerated? *Journal of Health Economics* 31 (1): 197-206.

Duval, S., and R. Tweedie. 2000. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 56 (2): 455-463.

Griffith, R., M. O'Connell, and K. Smith. 2019. Tax design in the alcohol market. *Journal of Public Economics* 172 (1): 20-35.

Hanson, A., and R. Sullivan. 2016. Incidence and salience of alcohol taxes: Do consumers overreact? *Public Finance Review* 44 (3): 344-369.

Hedges, L.V. 1983. A random effects model for effect sizes. *Psychological Bulletin* 93 (2): 388-395.

Hedges, L.V., and I. Olkin. 1985. *Statistical Methods for Meta-Analysis*. New York: Academic Press.

Hedges, L.V., and J.L. Vevea. 1998. Fixed- and random-effects models in meta-analysis. *Psychological Methods* 3 (4): 486-504.

Higgins, J.P.T., and S.G. Thompson. 2002. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 21 (11): 1539-1558.

Higgins, J.P.T., S.G. Thompson, J.J. Deeks, and D.G. Altman. 2003. Measuring inconsistency in meta-analysis. *BMJ (British Medical Journal)*: 327 (6 Sep.): 557-560.

Higgins, J.P.T., S.G. Thompson, and D.J. Spiegelhalter. 2009. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society A* 172 (1): 137-159.

IntHout, J., J.P.A. Ioannidis, M.M. Rovers, and J.J. Goeman. 2016. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 6 (7): art e010247.

Ioannidis, J.P.A., T.D. Stanley, and H. Doucouliagos. 2017. The power of bias in economics research. *The Economic Journal* 127 (605): f236-f265.

Johnston, R.B., J. Rolfe, R.S. Rosenberger, and R. Brouwer (eds.). 2015. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners.* New York: Springer.

Kenkel, D.S. 2005. Are alcohol tax hikes fully passed through to prices? Evidence from Alaska. *American Economic Review* 95 (2): 273-277.

Konstantopoulos, S., and L.V. Hedges. 2019. Statistically analyzing effect sizes: Fixed- and random-effects models. In H. Cooper, L.V. Hedges, and J.C. Valentine (eds.), T*he Handbook of Research Synthesis and Meta-Analysis,* 3rd edn.: 245-279. New York: Russell Sage.

Lin, L., and H. Chu. 2018. Quantifying publication bias in meta-analysis. *Biometrics* 74 (3): 785-794.

Lindhjem, H., and S. Navrud. 2008. How reliable are meta-analyses for international benefit transfers? *Ecological Economics* 66 (2-3): 425-435.

Nelson. J.P. 2015. Meta-analysis: Statistical methods. In Johnston, R.B., J. Rolfe, R.S. Rosenberger, and R. Brouwer (eds.). *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*: 329-356. New York: Springer.

Nelson, J.P., and P.E. Kennedy. 2009. The use (and abuse) of meta-analysis in environmental and natural resource economics: An assessment. *Environmental & Resource Economics* 42 (3): 345-377.

Nelson, J.P., and J.R. Moran. 2019. Effects of alcohol taxation on prices: A systematic review and meta-analysis of pass-through rates. *The B.E. Journal of Economic Analysis and Policy* (Ahead of Print, on-line 12/18/19). https://doi.org/10.1515/bejeap-2019-0134

Niskanen, W.A. 1962. The demand for alcoholic beverages: An experiment in econometric method. Rand Report P-2583. Santa Monica, CA: Rand Corp. https://www.rand.org/.

Paldam M. 2015. Meta-analysis in a nutshell: Techniques and general findings. *Economics: The Open-Access, Open-Assessment E-Journal* 9 (2015-11): 1-14.

Reed, W.R. 2015. A Monte Carlo analysis of alternative meta-analysis estimators in the presence of publication bias. *Economics: The Open-Access, Open-Assessment E-Journal* 9 (2015-30): 1-40.

Rhodes, W. 2012. Meta-analysis: An introduction using regression models. *Evaluation Review* 36 (1): 24-71.

Ringquist, E.J. 2013. *Meta-Analysis for Public Policy Management*. San Francisco: Jossey-Bass.

Shang, C., A. Ngo, and F.J. Chaloupka. 2018. The pass-through of alcohol taxes to prices in OECD countries. Discussion paper. Norman, OK: University of Oklahoma. https://www.researchgate.net

Stanley, T.D. 2008. Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics* 70 (1): 103-127.

Stanley, T.D., and H. Doucouliagos. 2012. *Meta-Regression Analysis in Economics and Business*. London, UK: Routledge.

Stanley, T.D., and H. Doucouliagos. 2015. Neither fixed nor random: Weighted least squares meta-analysis. *Statistics in Medicine* 34 (13): 2116-2127.

Stanley, T.D., S.B. Jarrell, and C. Doucouliagos. 2010. Could it be better to discard 90% of the data? A statistical paradox. *American Statistician* 64 (1): 70-77.

Stern, N. 1987. The effects of taxation, price control and government contracts in oligopoly and monopolistic competition. *Journal of Public Economics* 32 (2): 133-158.

Thompson, S.G., and J.P.T. Higgins. 2002. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* 21 (11): 1589-1573.

Thompson. S.G., and S.J. Sharp. 1999. Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine* 18 (20): 2693-2708.

Vevea, J.L., K. Coburn, and A. Sutton. 2019. Publication bias. In H. Cooper, L.V. Hedges, and J.C. Valentine (eds.), T*he Handbook of Research Synthesis and Meta-Analysis*, 3rd edn.: 383-429. New York: Russell Sage.

Viscusi, W.K. 2018. Best estimate selection bias in the value of a statistical life. *Journal of Benefit-Cost Analysis* 9 (2): 205-246.

Weyl, E.G., and M. Fabinger. 2013. Pass-through as an economic tool: Principles of incidence under imperfect competition. *Journal of Political Economy* 121 (3): 528-583.

Young, D.J., and A. Bielinska-Kwapisz. 2002. Alcohol taxes and beverage prices. *National Tax Journal* 55 (1): 57-73.

Appendix A: Pass-through studies in the review and meta-analysis

Ally, A.K., Y. Meng, R. Chakraborty, P.W. Dobson, et al. 2014. Alcohol tax pass-through across the product and price range: Do retailers treat cheap alcohol differently? *Addiction* 109 (12): 1994-2002.

Ardalan, A., and S.G. Kessing. 2017. Tax pass-through in the European beer market. Paper No. 47-2017. Marburg, DE: Philipps University of Marburg. https://www.econstor.eu/bitstream/10419/174343/1/47-2017_ardalan.pdf

Baker, P., and V. Brechling. 1992. The impact of excise duty changes on retail prices in the UK. *Fiscal Studies* 13 (2): 48-65.

Bako, B., and Z. Berezvai. 2013. Excise tax overshifting in the Hungarian beer market. Discussion paper. Budapest, HU: Corvinus University of Budapest. http://unipub.lib.uni-corvinus.hu/1263/1/tax_overflow.pdf

Barzel, Y. 1976. An alternative approach to the analysis of taxation. *Journal of Political Economy* 84 (6): 1177-1197.

Bergman, U.M., and N.L. Hansen. 2017. Are excise taxes on beverages fully passed through to prices? The Danish experience." Discussion paper. Copenhagen, DK: University of Copenhagen. http://web.econ.ku.dk/okombe/BergmanHansen.pdf

Carbonnier, C. 2013. Pass-through of per unit and ad valorem consumption taxes: evidence from alcoholic beverages in France. *B.E. Journal of Economic Analysis and Policy* 13 (2): 837-863.

Chua, J.D. 2000. *Essays in Public Economics*. Ph.D. dissertation. Cambridge, MA: Harvard University. http://www.worldcat.org/title/essays-in-public-economics/oclc/77070465

Conlon C.T., and N.S. Rao. 2016. Discrete prices and the incidence and efficiency of excise taxes. Discussion paper. New York, NY: New York University. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2813016

Cook, P.J. 1981. The effect of liquor taxes on drinking, cirrhosis, and auto accidents. In: Moore, M.H. and D.R. Gerstein (eds.), *Alcohol and Public Policy: Beyond the Shadow of Prohibition*: 255-285. Washington, DC: National Academy Press.

Hanson, A., and R. Sullivan. 2016. Incidence and salience of alcohol taxes: Do consumers overreact? *Public Finance Review* 44 (3): 344-369.

Harding, M., E. Leibtag, and M. Lovenheim. 2010. The heterogeneous geographic and socioeconomic incidence of cigarette and beer taxes: Evidence from Nielsen Homescan data." Discussion paper. Ithaca, NY: Cornell University. http://www.cemmap.ac.uk/forms/scanner_data/sd10_harding.pdf

Hindriks, J., and V. Serse. 2018. Heterogeneity in the tax pass-through to spirits retail prices: Evidence from Belgium. Paper 2018/04. Louvain, BE: Center for Operations Research and Econometrics, Catholic University of Louvain. https://dial.uclouvain.be/pr/boreal/object/boreal:195857

Hunt, P., L. Rabinovich, and B. Baumberg. 2010. Preliminary assessment of the economic impacts of alcohol pricing policy options in the UK. Rand Europe Report to the Home Office TR-858-HO. Brussels, BE: Rand Europe. https://www.rand.org

Hunt, P., R. Smart, and R.L. Pacula. 2018. Are alcohol taxes still an effective policy lever for raising alcohol prices to improve health? Evidence from the Great Recession. Discussion paper. Santa Monica, CA: Rand Corp.

Kenkel, D.S. 2005. Are alcohol tax hikes fully passed through to prices? Evidence from Alaska. *American Economic Review* 95 (2): 273-277.

Loretz, S., and G. Zwirn. 2015. Taxation of alcoholic beverages: Estimation of tax revenue effects in the Netherlands and Europe. Research report. Vienna, AT: Institute for Advanced Studies. http://irihs.ihs.ac.at/3604/

Mathes, M.T., and C.S. Carpenter. 2015. New evidence on beer prices and beer taxes. In: Mathes, M.T., *Essays on Health Economics: Effect of Economic Forces on Drinking and Smoking-Related Outcomes*: 64-82. Ph.D. dissertation. Nashville, TN: Vanderbilt University. https://search.proquest.com/docview/1724663795

Niskanen, W.A. 1962. The demand for alcoholic beverages: an experiment in econometric method. Rand Report P-2583. Santa Monica, CA: Rand Corp. https://www.rand.org/content/dam/rand/pubs/papers/2008/P2583.pdf

Parsons, C. 2007. Towards an empirical industrial organization analysis of the Japanese beer industry." Paper CITS WP 2007-01. Yokohama, JP: Yokohama National University. https://www.researchgate.net

Pipoblabanan, C. 2008. *Essays on Parametric and Nonparametric Estimation of Market Structure and Tax Incidence in the U.S. Brewing Industry*. Ph.D. dissertation. Corvallis, OR: Oregon State University. https://ir.library.oregonstate.edu/concern/graduate_thesis_or_dissertations/x059cb29m

Rabinovich, L., P. Hunt, L. Staetsky, S. Goshev, et al. 2012. Further study on the affordability of alcoholic beverages in the EU: A focus on excise duty pass-through, on- and off-trade sales, price promotions and statutory regulations. Report for the European Commission. Cambridge, UK: Rand Europe. https://www.rand.org/pubs/technical_reports/TR1203.html

Russell, C., and C. Van Walbeek. 2016. How does a change in the excise tax on beer impact beer retail prices in South Africa? *South African Journal of Economics* 84 (4): 555-573.

Shang, C., A. Ngo, and F.J. Chaloupka. 2018. The pass-through of alcohol taxes to prices in OECD countries. Discussion paper. Norman, OK: University of Oklahoma. https://www.researchgate.net

Shrestha, V., and S. Markowitz. 2016. The pass-through of beer taxes to prices: Evidence from state and federal tax changes. *Economic Inquiry* 54 (4): 1946-1962.

Siegel, M., J. Grundman, W. DeJong, T.S. Naimi, et al. 2013. State-specific liquor excise taxes and retail prices in eight U.S. states, 2012. *Substance Abuse* 34 (4): 415-421.

Stehr, M. 2007. The effect of Sunday sales bans and excise taxes on drinking and cross-border shopping for alcoholic beverages. *National Tax Journal* 60 (1): 85-105.

Tasarika, E. 2001. *Aspects of International Taxation*. Ph.D. thesis. Exeter, UK: University of Exeter. https://www.researchgate.net/profile/Euamporn_Phijaisanit/research

Tiwary, R. 2011. *Determinants of Tax Pass-Through Rates: A Study of the U.S. Beer Industry*. Master's thesis. Amherst, MA: University of Massachusetts Amherst. https://scholarworks.umass.edu/resec_theses/

Young, D.J., and A. Bielinska-Kwapisz. 2002. Alcohol taxes and beverage prices. *National Tax Journal* 55 (1): 57-73.