

The relative effectiveness of R&D tax credits and R&D subsidies: A comparative Meta-Regression Analysis

Dimos, C., Pugh, G., Hisarciklilar, M., Talam, E. & Jackson, I
Author post-print (accepted) deposited by Coventry University's Repository

Original citation & hyperlink:

Dimos, C, Pugh, G, Hisarciklilar, M, Talam, E & Jackson, I 2022, 'The relative effectiveness of R&D tax credits and R&D subsidies: A comparative Meta-Regression Analysis', *Technovation*, vol. 115, 102450.

<https://doi.org/10.1016/j.technovation.2021.102450>

DOI 10.1016/j.technovation.2021.102450

ISSN 0166-4972

ESSN 1879-2383

Publisher: Elsevier

© 2022, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

The relative effectiveness of R&D tax credits and R&D subsidies: A comparative Meta-Regression Analysis

Abstract

There are large primary literatures that evaluate the effectiveness of either R&D tax credits or R&D subsidies in promoting private R&D. However, this Meta-Regression Analysis, by investigating these literatures jointly, is the first study that systematically measures and compares the effectiveness of these two policy instruments. After controlling for publication selection and sources of heterogeneity, we find that both tax credits and subsidies induce additional private R&D and that neither instrument systematically outperforms the other. However, whereas subsidy effects are increasing over time tax credit effects are not. Although their respective effects are “small”, they are not negligible: in round terms, an additional \$1 of public R&D support of either type induces 7.5 cents of additional private R&D expenditure. Sources of heterogeneity in the reported effects include: tax credits are most effectively delivered as “incremental” schemes, are more effective in economies with a balanced “policy-mix” regime, and are generally less effective for micro firms and SMEs than for large firms; while subsidies are more effective for manufacturing firms, although not for high-tech firms, and are more effective than tax credits in economies predominantly using subsidies. Finally, we argue for the importance of statistical power in the design of evaluation studies.

Keywords: R&D tax credits; R&D subsidies; Meta-regression analysis; Publication bias; Policy evaluation; Additionality

JEL Classification: C10; H23; H25; H59; O31; O38

1. Introduction

Causal relationships between research and development (R&D), technical progress and sustained per capita income growth have long been received wisdom (Schumpeter, 1942; Solow, 1956). However, due to market failures, the socially optimal level of R&D investment is not realised (Nelson, 1959; Arrow, 1962; Usher, 1964). Accordingly, public authorities have adopted a range of public R&D support instruments, in particular indirect fiscal R&D support (R&D tax credits) and direct R&D support (R&D subsidies), which are among the ‘main innovation policy levers’ and in the short run ‘seem the most effective’ (Bloom et al., 2019: 163 and 180). Yet, neither theory nor the available empirical evidence arbitrates on their relative effectiveness. Accordingly, the contribution of this study is to measure and compare the effectiveness of these two policy instruments. There are extensive literatures that consider separately the effectiveness of R&D tax credits and R&D subsidies, and that identify sources of heterogeneity in their effects. Nevertheless, as noted by Busom et al. (2014: 572) ‘an explicit and comparative analysis of both tools remains to be done’ (see also Becker, 2015: 925).

Recently, two studies have used Meta-Regression Analysis (MRA) (Castellacci and Lie, 2015; Dimos and Pugh, 2016) to investigate, respectively, the R&D tax credit and R&D subsidy literatures, while Becker (2015) and Bloom et al. (2019) have provided narrative reviews of the empirical evidence from both literatures. By design, the MRA studies of Castellacci and Lie (2015) and Dimos and Pugh (2016) study the effectiveness of single instruments in isolation. This study is different, because it is a comparative MRA, designed to compare the relative effectiveness of the two instruments. Unfortunately, methodological differences between the two previous MRA studies – in particular, their approach to estimating representative empirical effects from their respective literatures – preclude using them to compare the effectiveness of the two instruments. (For a comparison between the previous MRA studies and the current study, see Appendix A.) In contrast, this study applies MRA to both literatures jointly to determine the relative effectiveness of R&D tax credits and R&D subsidies after taking account of

- (a) the heterogeneity of samples and methodologies in each literature and
- (b) the degree – if any – of publication selection bias (henceforth, publication bias) in each literature.

In the absence of either theory or a body of empirical evidence on the relative effectiveness of R&D tax credits and R&D subsidies, the research in this paper is exploratory rather than theory testing.

We find that both tax credits and subsidies give rise to additionality of a similarly small size and that, as a corollary (Ioannidis, et al. 2017), given typical sample sizes, both literatures suffer from low statistical power (the statistical power of the two literatures is explored in Appendix E). Moreover, mindful of the warning by David et al. (2000: 500) against ‘broad empirical generalisation’ of R&D support effects, we identify sources of heterogeneous findings both between and within these two literatures. Either absence or presence of

heterogeneous effects across types of firm, industry and/or country is informative for policy makers, suggesting whether R&D policies are broadly transferable or context-specific.

The paper is structured as follows. Section 2 provides context and theoretical background on the two policy instruments. Section 3 discusses the MRA database. Sections 4 and 5 present our MRA models and our main results. Section 6 extends our MRA to a subset of studies that report effect sizes as elasticities, thus measuring in value terms the responsiveness of R&D expenditure to R&D support. Section 7 concludes.

2. Context and theoretical background

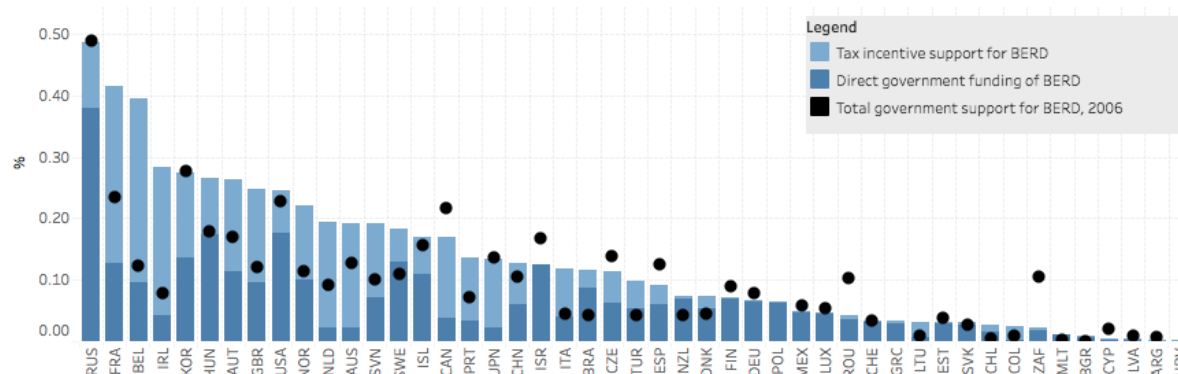
Although R&D tax credits and R&D subsidies are not perfect substitutes, they are sufficiently similar for their effects to be meaningfully compared by meta-regression analysis. Narrative reviews have compared the effects of a wide range of public innovation policies including these two instruments (Becker, 2015; Bloom et al., 2019) to provide evidence on ‘the best use of resources’ (Bloom et al., 2019: 178). Given that R&D tax credits and R&D subsidies have the common objective of inducing additional private R&D investment by reducing its cost (David et al., 2000; Bloom et al., 2019), policy makers have an interest in their relative effectiveness. Indeed, these two policy instruments are typically considered together (OECD, 2018 – see Figure 1 below; Bloom et al., 2019) and are often used together by firms to support their R&D (Bérubé and Mohnen, 2009; Castellacci and Lie, 2015).

Contrary to tax credits, which are provided *ex post*, subsidies are typically provided *ex ante* or during the private R&D investment. Moreover, tax credits are available to all eligible firms by taking into account either their overall R&D spending (volume-based system) or their excess R&D spending above specified thresholds (incremental-based system), whereas subsidies target specific projects with high social returns and their selection for funding rests largely upon the information available to and the discretion of the public agency. Tax credits are relatively immune to policy inefficiencies, since they are based upon firms’ optimisation decisions. In contrast, subsidies require a bureaucratically intensive selection process and are more susceptible to policy inefficiencies, due to information asymmetries between recipient firms and programme managers and – according to public choice theory – the potentially self-interested objectives of programme managers (Dimos and Pugh, 2016: 798).

Both R&D tax credits and R&D subsidies are widely used to promote private R&D investment. Figure 1 shows government spending in 2016 on R&D tax incentives (mainly tax credits) and direct R&D funding (mainly subsidies) expressed as a percentage of GDP together with total government support in 2006. While both overall public R&D support and the relative size of spending on the two measures vary across countries, in most cases governments use both instruments.¹

¹ The primary literature investigated by the current study is restricted to tax credits, the main type of expenditure-based tax incentives, and subsidies encompassing grants and/or (low-interest) loans. For other forms of R&D support, see OECD (2015) and Becker (2015).

Figure 1. Direct government funding and tax support for business R&D, 2016 (as a percentage of GDP)



* BERD (Business Expenditure on R&D)

Note: Data on tax incentive support not available for Israel and Malta.

Source: OECD R&D tax incentive database (November 2018). Accessed from: <http://www.oecd.org/sti/rd-tax-stats.htm>

In 2015, 28 of the 34 OECD countries and a number of non-OECD economies provided R&D expenditure-based tax incentives (OECD, 2015). From 2000 to the onset of the Global Financial Crisis (GFC), several OECD countries increased their reliance on R&D tax incentives to promote R&D investment. Tax credits are cheaper to manage and are the more market-conforming approach, being rights-based (applying equally to all eligible firms), led by firms' decision making and thus entailing minimal governmental discretion. However, although tax credits are less prone to political and institutional instability, this mode of support also proved the more vulnerable to market instability. Reflecting dependence on profits, the relative importance of tax incentives declined briefly in the aftermath of the GFC (OECD, 2015) while governments tended to maintain direct funding to mitigate the impacts of the crisis on business R&D (Hud and Hussinger, 2015). Direct funding is also consistent with a renewed interest in industrial policy (Stiglitz and Greenwald, 2015: 20-24), given that subsidies can be used to support R&D projects according to governmental perceptions of their social rate of return, targeting types of R&D judged to be particularly undersupplied. For example, whereas complex tax regulations are often held to bias the use of tax credits towards large firms and away from firms in traditional manufacturing industries, subsidies may be used to redress the balance.

The literature provides limited guidance on the relative effectiveness of tax credits and subsidies in different contexts. In particular, the two instruments do not reach and/or affect all firms and sectors equally. Peneder (2008) argues that access to finance is inversely related to firm size and Busom et al. (2014) highlight the mitigation of financial constraints as a source of differential support effectiveness according to firm size (as subsidies providing ex ante help may be more appropriate for financially constrained SMEs while ex post tax credits may be more appropriate for large firms with greater financial resources). Accordingly, there is no expectation of a uniformly preferred approach for public support across firms of different size. According to Castellacci and Lie (2015: 819), different sectors exhibit varying degrees of 'market competition, technological opportunities' and 'intensity of knowledge diffusion and spillover effects', which condition the way in which firms organise their innovative

activities and thus give rise to heterogeneous responses to R&D incentives. For example, Yang et al. (2012) associate more (less) fertile technological environments with high-tech (traditional) industries and thus potentially different responses to public R&D support.

At the firm level, the two types of support give rise to different ranges of potential effects on private R&D investment. Tax credits delivered at arm's length after private R&D investment has taken place are not expected to give rise to crowding-out effects (David et al., 2000). Hence, tax credits may: either be a deadweight loss (i.e. no effect) – because the R&D investment might have gone ahead anyway, the subsequent cost of the tax credit to public funds may not yield any corresponding public benefit (Baghana and Mohnen, 2009; Mohnen, n.d.); or yield additionality, because, according to Duguet (2012: 408), 'a firm can integrate the tax credit into its investment decision process and decide to invest more because the deduction exists'. In contrast, subsidy effects range from crowding out (the subsidy substituting for private investment) through no effect (the subsidy is merely added to unchanged private investment) to additionality (the subsidy induces increased private R&D investment) (Dimos and Pugh, 2016). Notwithstanding, this asymmetry of potential outcomes does not imply that additionality is more likely from one instrument than from the other.

3. MRA database and preliminary investigation

3.1 Combining the two literatures

The validity of any meta-regression analysis depends on the coherence of the effect sizes extracted from the primary literature (Stanley and Doucouliagos, 2012: 13-15). In the case of the present study, because we are comparing across two related literatures, the coherence of the effect sizes requires further consideration. To ensure comparable effect sizes, our sample selection principle is the same across both the tax credit and the subsidy literatures: i.e. to compile *all* estimates of the rate at which private R&D expenditure changes due to either tax credits or subsidies. This sampling principle ensures quantitative comparability of the effectiveness of the two instruments. Moreover, our empirical design accounts for both within- and between-instrument heterogeneities. We use study-level fixed effects as well as “moderator” variables to control for both within-instrument heterogeneities (e.g. “incremental” versus “total” tax-credit schemes) and between-instrument heterogeneities, which include differing incidence of effectiveness (e.g. with respect to high-tech firms, SMEs and industrial categories) and potentially differing applicability between sectors (given that neither tax credits nor subsidies are necessarily applied uniformly across either firm sizes or sectors). For a complete list of moderator variables, see Table 1 in Section 4 below. In sum, a common effect size and analysis conditional on a wide range of moderator variables allow valid quantitative comparison of the effects of R&D tax credits and R&D subsidies on private R&D expenditure.

We searched the EconLit online database as well as making Google-Scholar searches during April 2018 using keywords including “R&D tax credits (subsidies)”, “R&D tax credit (subsidy) effect”, “R&D tax incentives” and “R&D tax incentives effect”. In addition, we

checked the references from the identified studies; in particular, the references in meta-regression analyses conducted on the tax credit and subsidy literatures separately (Castellacci and Lie, 2015; Dimos and Pugh, 2016). We then coded all studies written in English, published in or after 2000, and reporting additionality-ratio effects. To investigate the relative effectiveness of R&D tax credits and R&D subsidies, we require comparable effect sizes across the two literatures. Hence, our sample is restricted to studies estimating the rate at which private R&D expenditure (i.e. net of tax credit or subsidy amount) changes due to, respectively, tax credits or subsidies, both of which may be measured as values or captured by dummy variables.^{2,3}

The resulting MRA database comprises 251 tax credit effects from 12 studies and 347 subsidy effects from 25 studies. (These studies are listed in Appendix B.) In our database, the typical tax credit (subsidy) study reports 21 (14) effect sizes with the median being 15.5 (14); across these studies, the minimum number is 1 (1) and the maximum 72 (34).

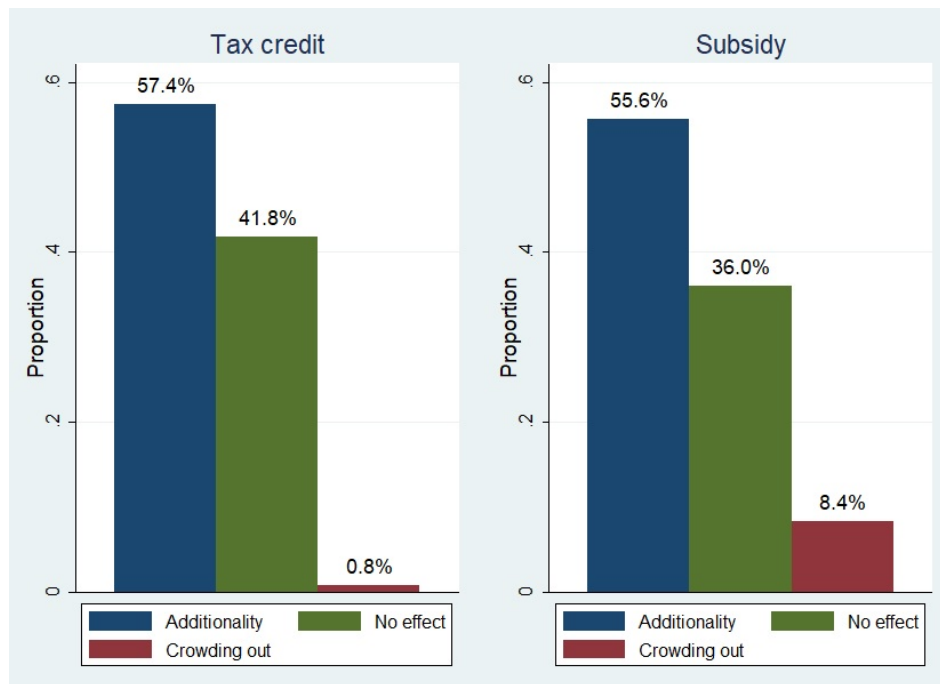
3.2 Effect sizes and public support outcomes

Figure 2 depicts the reported effects from both literatures. In accord with theory, the subsidy literature reveals all three possible outcomes: 8.4% (29 effects) of the reported estimates indicate crowding out, while additionality and no-effect outcomes appear in proportions similar to the tax credit literature. In the tax credit literature almost six out of ten estimates correspond to additionality while the rest suggest no effect. Although we do not expect to find crowding-out effects in the tax credit literature, two statistically significant negative effects (0.8%) are reported by Lee (2011) and Ho (2006) respectively. Both authors comment on these apparently perverse results, with Ho (2006) reflecting on the possibility that the comparison group might not have been valid. However, we would add sampling error as an explanation. Indeed, from the meta-perspective of the total number of estimates in the literature, we might expect more such perverse results to be reported (Stanley and Doucouliagos, 2012: 56).

² The effects of both types of public support may be estimated within either a matching or a regression framework. Both yield three possible effects: positive and statistically significant – additionality; statistically insignificant – absence of a tax credit/subsidy effect; and negative and statistically significant – crowding out. As explained in Section 2, in micro-level studies we do not expect to find crowding-out effects in the tax credit literature.

³ Although there exist a few recent studies on the effect of R&D subsidies, none of these reports effect sizes on private R&D expenditure and, thus, cannot be added to our database.

Figure 2. Reported effects from public support of private R&D



Source: authors.

3.3 Transformation of reported effects into PCCs

In the MRA database, roughly equal proportions of effect sizes arise from parametric (regression) and non-parametric (matching) approaches, and mostly have a measurement-unit-dependent interpretation. To make these effects comparable, we follow Doucouliagos and Stanley (2009) and Stanley and Doucouliagos (2012) by transforming these effects into partial correlation coefficients (PCCs). This procedure provides a unit-free measure of the magnitude and direction of the association between two variables; in our case, between the R&D tax credit and private R&D expenditure or between the R&D subsidy and private R&D expenditure, holding other variables constant. The PCC and its standard error (SE_{PCC}) are calculated as follows:

$$PCC = \frac{t}{\sqrt{t^2 + df}} \quad (1)$$

$$SE_{PCC} = \sqrt{\frac{1 - PCC^2}{df}} \quad (2)$$

where t stands for the t-statistic on the estimated tax credit or subsidy effect and df for the degrees of freedom reported in the primary studies.

4. Meta-Regression Analysis

In this Section, we investigate sources of heterogeneity in reported effect sizes both within and between the two literatures. In turn, we estimate “authentic” empirical effects – following common terminology in MRA – “beyond” publication selection bias and sources of heterogeneity, where these may be sample characteristics – e.g. firm, industry and/or country – or different research practices.

MRA estimates authentic empirical effects “beyond” – i.e. after controlling for or filtering out – publication bias (Stanley, 2005), which arises from individual decisions to report estimates only of a certain “theory-consistent” sign and/or of a sufficient size to compensate for low precision, thereby yielding “significant” estimates. Conversely, low-precision estimates of the “wrong” sign, or estimates too small to offset high standard errors, are not reported. Both quantitative investigation of estimates reported in major economics journals (Brodeur et al., 2016) and survey responses from academic economists (Necker, 2014) suggest that researchers are incentivised to ‘search for specifications delivering just-significant results and ignore specifications giving just-insignificant results in order to increase their chances of being published’ (Brodeur et al., 2016: 2). Within the field of R&D studies, Klette et al. (2000: 487) warned against ‘a publication filter, self-imposed by researchers, or imposed by editors and referees considering non-significant coefficients to be of little interest’. The aggregate outcome of these individual decisions shaped by such a “filtered” publication process can skew the distribution of effects reported in empirical literatures. In turn, the mean effect across a literature reflects not only the true effect but also publication selection bias. The corresponding role of MRA is both to identify publication bias, which is endemic in empirical literatures in economics (Stanley and Doucouliagos, 2012: 52), and to estimate authentic effect sizes in empirical literatures controlling for publication bias (Stanley, 2005).

When publication selection is absent, effect sizes are independent of their standard errors, from which follows a simple model for estimating the authentic effect from an econometric literature while controlling for publication bias (Stanley, 2005; Stanley, 2008; and Stanley and Doucouliagos, 2012):

$$PCC_i = \beta_0 + \beta_1 SE_i^{PCC} + \varepsilon_i \quad (3)$$

where $i = 1, \dots, n$ indexes the n individual estimates reported in the primary literature, SE_i^{PCC} denotes the standard error of the i^{th} PCC and ε_i is the usual regression error. The statistical significance of β_1 indicates the presence of publication bias and its sign the direction, while rejection of the null hypothesis $\beta_0 = 0$ is evidence of an authentic effect “beyond” publication selection bias and the magnitude of this coefficient is an estimate of the authentic effect in terms of the PCC (Stanley, 2005).

Weighted Least Squares (WLS) estimation is employed to correct for heteroskedasticity, which is a characteristic of Eq. (3), because the variance of the PCC_i (and, thus, the variance of ε_i) is not constant. WLS estimation is implemented by dividing Eq. (3) by the standard error of PCC_i (SE_i^{PCC}) (Stanley and Doucouliagos, 2012: 61), which not only addresses

heteroskedasticity but also implements precision weighting (i.e. more precise estimates are given greater weight):

$$\frac{PCC_i}{SE_i^{PCC}} = t_i = \beta_1 + \beta_0 \left(\frac{1}{SE_i^{PCC}} \right) + v_i \quad (4)$$

where t_i is the t-statistic on each PCC (i.e. the t-statistic on each corresponding effect reported in the primary literature), $v_i = \varepsilon_i / SE_i^{PCC}$ is the error term corrected for heteroskedasticity, and $1/SE_i^{PCC}$, the inverse of the SE on the PCC , is the precision term.⁴ In the transformation of Eq. (3) into Eq. (4) the coefficients β_0 and β_1 change place but their interpretation is unchanged.

Because each study reports a multitude of effect sizes, studies reporting a relatively large number of effect sizes exert undue influence on both descriptive statistics and our later regression estimates. To offset this undue influence, we also weight each estimated effect by the inverse number of estimates reported in the source study.⁵ Hence, our estimates reflect two types of weighting: (i) precision weighting; and (ii) study weighting.

Our approach is not to estimate tax credit and subsidy models separately, by dividing our MRA database into its component literatures, but instead to pool the two samples and use the full resources of our data to exploit efficiency gains. This procedure is explained in a standard reference work for applied economics (Greene, 1993: 236) and has also been recommended for political science (Brambor et al., 2006: 78):

Some scholars test conditional hypotheses by splitting their data into categories (such as male and female or north and south) across which the effect of some independent variable X is supposed to differ. Instead of using explicit interaction terms, they simply run separate regressions on each of these categories. While this is a perfectly reasonable way to test conditional hypotheses, there is no real gain in terms of interpretability and there will always be a loss of efficiency due to the smaller sample sizes.

For the fully specified WLS model set out in Eq.8 and estimated on the full sample (see Table 2a in Section 5), we tested the null of variance equality between the tax credit and subsidy subsamples, mindful that rejection suggests using the subsamples separately (Greene, 1993: 236). When we use the full sample, the null hypothesis of variance equality is rejected ($p=0.000$). However, after using standard approaches for identifying outliers, we discovered that one study in the subsidy literature (i.e. Ugur et al., 2015) has such different statistical characteristics from the rest of the sample that it alone is responsible for variance inequality across the two components of the full sample.⁶ Upon excluding Ugur et al. (2015) from the sample, we cannot reject the null of variance equality ($p=.1830$). Accordingly, we estimate our WLS model first using the full sample excluding Ugur et al. (2015) and second – as a

⁴ Following Fisher (1954:194), Eq. (1) divided by Eq. (2) yields t , which is the t-statistic on the effects reported in the primary literature.

⁵ The study-weighted PCCs are close to the study-unweighted estimates, which suggests the absence of severe bias from studies reporting relatively large numbers of estimates (see Table 1 below).

⁶ For example, the mean and standard deviation of the PCC of the full sample without Ugur et al. (2015) are respectively: .0676 and .0991. The contrasting values for Ugur et al. (2015) alone are: -.0153 and .0368.

robustness check – using separate sample estimation including Ugur et al. (2015). No further data cleaning was necessary.

To estimate this model for the tax credit and the subsidy literatures jointly, Eq. (5) augments Eq. (4) with an intercept shift dummy for the tax credit literature (D_i^{tax}) – 1 (0) denotes estimates from the tax credit (subsidy) literature – and an interaction term between D_i^{tax} and $1/SE_i^{PCC}$. In this specification, subsidy effects are estimated directly but tax credit effects are derived post estimation as linear combinations of estimated coefficients.⁷ This yields the same point estimates as separate sub-sample regression but with slightly smaller standard errors:

$$t_i = \beta_1 + \beta_0 \left(\frac{1}{SE_i^{PCC}} \right) + \beta_2 D_i^{tax} + \beta_3 \left(\frac{1}{SE_i^{PCC}} D_i^{tax} \right) + v_i \quad (5)$$

where β_1 (the intercept) controls for and measures publication bias in the subsidy literature; β_2 measures the effect of D_i^{tax} , which is the difference between publication bias in the tax credit literature and publication bias in the subsidy literature; and the sum of the estimated coefficients $\beta_1 + \beta_2$ measures the publication bias in the tax credit literature. Accordingly, having controlled for publication bias, β_0 measures the authentic effect for the subsidy literature, β_3 measures the difference between the authentic effect for the tax credit literature and the authentic effect for the subsidy literature, and the sum of estimated coefficients $\beta_0 + \beta_3$ measures the authentic effect for the tax credit literature.

We model sources of heterogeneity with “moderator variables” Z_i (see Table 1). These are of two kinds, both of which condition effect sizes: (i) “internal” moderators, which are dummy variables that capture characteristics of the estimates reported in the primary literature (e.g. sample and research characteristics); and (ii) “external” moderators from outside the primary literature (in the present case, to control for differing preferences across countries regarding policy mix).⁸ Both types of moderator variables are specified in the same way in our MRA. Accordingly, Eq. (6) augments Eq. (5) with two sets of variables:

- (a) moderator variables interacted with the precision effect, i.e. Z_{mi}/SE_i^{PCC} ; and
- (b) moderator variables interacted with the precision effect and the tax credit dummy, i.e. $(Z_{mi}/SE_i^{PCC})D_i^{tax}$, to allow for differential moderator effects between the two literatures.

⁷ These linear combinations and their standard errors are obtained by Stata’s *lincom* command. Illustrative examples of the linear combinations estimated below are given in Appendix C.

⁸ We are thankful for an anonymous suggestion that some countries traditionally rely mostly on direct public incentives for R&D, others rely most on indirect public incentives, and there is still a third group that combines the use of direct and indirect public incentives. Hence, independently of the publication bias, national effects may play a role in explaining eventual differences in the size of the effect identified. Accordingly, we augment our “internal” moderators with two “external” moderators; namely, dummy variables for: (i) estimates obtained from samples from countries and periods with tax credits accounting for at least 60 per cent of total public R&D support; and (ii) estimates obtained from samples from countries and periods with subsidies accounting for at least 60 per cent of total public R&D support. The omitted category comprises estimates obtained from samples from countries and periods with both tax credits and subsidies accounting for between 40 and 60 per cent of total public R&D support. The data sources on the relative incidence of direct (subsidy) and indirect (tax credit) support for R&D used to construct these variables were OECD (2015) together with the 2017 and 2018 country reports on R&D tax incentives available via OECD (2018).

$$t_i = \beta_1 + \beta_2 D_i^{tax} + \beta_0 \left(\frac{1}{SE_i^{PCC}} \right) + \beta_3 \left(\frac{1}{SE_i^{PCC}} D_i^{tax} \right) + \sum_m \beta_{m1} \left(\frac{Z_{mi}}{SE_i^{PCC}} \right) + \sum_m \beta_{m2} \left(\frac{Z_{mi}}{SE_i^{PCC}} D_i^{tax} \right) + v_i \quad (6)$$

where m indexes moderator variables. The estimates of β_{m1} measure the effect of each of the m moderator variables conditioning estimated effects in the subsidy literature; the estimates of β_{m2} measure the difference between each moderator effect in the tax credit and in the subsidy literatures; and the sum of each pair of estimated coefficients $\beta_{m1} + \beta_{m2}$ measures the respective moderating effects for the tax credit literature.

The authentic effect in the subsidy literature is given by the sum of β_0 , the estimated coefficient on the precision term ($1/SE_i^{PCC}$), and the β_{m1} , the estimated coefficients on each of the interacted moderator variables (Z_{mi}/SE_i^{PCC}) weighted by the study-weighted means of the corresponding moderator variables. Each study-weighted mean is the proportion of appearances of a moderator in the subsidy sample (so that a moderator appearing in 20% of the sample has twice the influence of one appearing in 10%) adjusted for whether it appears in studies reporting relatively few or relatively many estimates. Hence, a moderator appearing in 10% of the estimates in the respective primary literature but concentrated in studies reporting many estimates would be down weighted compared to a moderator appearing in 10% of the estimates but present in studies reporting relatively few estimates. (Therefore, the more studies in which a moderator appears the greater its weight).⁹ Similarly, the authentic effect in the tax credit literature is given by the sum of the precision effect and the combined effect of the moderator variables. This is calculated as the sum of two sums: β_0 and β_3 , the estimated coefficient on the differential precision term for the tax credit literature ($1/SE_i^{PCC}$) D_i^{tax} ; and the sum of each of the β_{m1} and each of the corresponding β_{m2} , the estimated coefficients on the differential tax credit moderator variables (Z_{mi}/SE_i^{PCC}) D_i^{tax} , weighted by the respective study-weighted means. (Appendix C gives indicative examples of the syntax used to implement these calculations.)

Our model specification takes into account the (unbalanced) panel structure of our data (Rosenberg and Loomis, 2000; Nelson and Kennedy, 2009), which arises because studies in the primary literature typically report multiple estimates. We address this in three ways: (i) by reporting weighted estimates giving each study equal influence; (ii) by reporting cluster-robust SEs, which are robust to arbitrary patterns of dependence among the residuals from estimates from the same study; and (iii) by augmenting Eq. (6) to control for the moderating effects of each specific study on the estimated authentic effects of R&D support:

$$t_{is} = \beta_1 + \beta_2 D_i^{tax} + \beta_0 \left(\frac{1}{SE_{is}^{PCC}} \right) + \beta_3 \left(D_i^{tax} \frac{1}{SE_{is}^{PCC}} \right) + \sum_m \beta_{m1} \left(\frac{Z_{mis}}{SE_{is}^{PCC}} \right) + \sum_m \beta_{m2} \left(D_i^{tax} \frac{Z_{mis}}{SE_{is}^{PCC}} \right) + \sum_s \left(\mu_s \frac{1}{SE_{is}^{PCC}} \right) + v_{is} \quad (7)$$

⁹ Study-weighted means for each moderator variable are obtained by using Stata's *summarize* command with *aweight* being the inverse number of estimates reported in each study.

where s indexes the studies in the primary literature. The study-specific effects are denoted by μ_s (fixed effects – i.e. a dummy variable for each study), which do not enter the equation *other than* by way of interaction with the inverse standard error of the PCC – i.e. $\mu_s(1/SE_{is}^{PCC})$.¹⁰

Our “internal” moderator variables include all those used in both the previous MRAs on the tax credit and subsidy literatures – respectively, Castellacci and Lie (2015) and Dimos and Pugh (2016) – together with additional ones suggested by Becker (2015).¹¹ These variables are listed together with study-unweighted and study-weighted descriptive statistics in Table 1. Stanley and Doucouliagos (2012: 85) distinguish between precision-interacted “ Z ” moderators directly influencing the authentic effect and intercept shift “ K ” moderators revealing sources of heterogeneity in the estimated publication bias. Castellacci and Lie (2015) include only Z moderators, whereas Dimos and Pugh (2016) also include K moderators. In this study, we specify our model with an extensive range of Z moderators, because we are above all interested in the heterogeneity of the authentic empirical effects between and within the tax credit and subsidy evaluation literatures. In addition, we include a single K moderator – “Year of publication” (*PubYear*) – to model the evolution of publication bias over time. Our objective is to measure and control for rather than to analyse publication bias.

Among our Z moderator variables is the “Starting point of data” (Table 1), which captures potential time variation in the authentic effects. Our K moderator allows the potential evolution of the authentic effect to be estimated conditional on the potential evolution of publication bias. We do not use the same variable to model time variation both in the authentic effects and in publication bias. Whereas the period covered by the data may influence the authentic effect but is not obviously related to publication bias, the year of publication should not influence the authentic effect but may capture current proclivities that influence publication practices.

Our fully specified model is set out in Eq. (8):

$$\begin{aligned} t_{is} = & \beta_1 + \beta_2 D_i^{tax} + \beta_0 \left(\frac{1}{SE_{is}^{PCC}} \right) + \beta_3 \left(D_i^{tax} \frac{1}{SE_{is}^{PCC}} \right) + \sum_m \beta_{m1} \left(\frac{Z_{mis}}{SE_{is}^{PCC}} \right) \\ & + \sum_m \beta_{m2} \left(D_i^{tax} \frac{Z_{mis}}{SE_{is}^{PCC}} \right) + \sum_s \left(\mu_s \frac{1}{SE_{is}^{PCC}} \right) + \gamma_1 PubYear_{is} \\ & + \gamma_2 (D_i^{tax} * PubYear_{is}) + v_{is} \end{aligned} \quad (8)$$

where γ_1 estimates the moderating effect of *PubYear* on publication bias in the subsidy literature, γ_2 measures the difference in the moderating effect of *PubYear* on publication bias between the two literatures, and $\gamma_1 + \gamma_2$ measures the moderating effect of *PubYear* on publication bias in the tax credit literature.

¹⁰ In the estimated models reported below, we include only those study-effects that prove to be statistically significant in a standard “general to specific” testing-down procedure (Stanley and Doucouliagos, 2012: 90).

¹¹ The only moderator variable that is not used from Castellacci and Lie (2015: 824) is “subsidy”, which captures whether ‘public R&D subsidies received by firms are included as control in the specification’. However, because only one study exists in which *all* firms receive subsidies in addition to tax credits, we cannot make valid inferences on the moderating effect of subsidy provision on tax credit effectiveness.

Although WLS is the standard, it is not the only approach to estimation used in MRA (Koetse et al., 2010; Stanley and Doucouliagos, 2013). Best practice is to check the robustness of MRA findings across different estimators (Stanley and Doucouliagos, 2012: 104; Stanley et al., 2013). To this end, we implement two more approaches to the estimation of our model. First, we use robust regression to address uncertainty regarding the inclusion of observations that may be outliers or/and exerting high leverage (i.e. exerting undue influence on the regression estimates).¹² Robust regression weighting precludes the additional use of study weighting. Second, we use Bayesian Model Averaging (BMA) to address uncertainty regarding the choice of moderator variables by providing a check on the relevance of each moderator variable across all possible combinations of the specified moderator variables (Iršová and Havránek, 2013, for a recent application to MRA).¹³ By construction, BMA estimation uses all specifications stemming from all possible combinations of the right-hand side variables to derive the estimates. Consequently, using the available software for the whole sample with interaction terms, we cannot avoid estimating misspecified models. For example, a model including a moderator variable ($\frac{Z_{mis}}{SE_{PCC}^{PCC}}$) without its corresponding interaction with the tax credit dummy ($D_i^{tax} \frac{Z_{mis}}{SE_{PCC}^{PCC}}$) does not yield the subsidy effect but an aggregate public support effect from both the tax credit and the subsidy literatures. Accordingly, we apply BMA to separate sample regressions.

In sum, we estimate Eq. (8) on the pooled sample by: (i) WLS (with study weights); and (ii) robust regression (without study weights – i.e. each observation is given equal weight). The corresponding results are reported in Table 2a (Section 5). In addition, we omit the interaction terms involving the tax credit dummy from Eq. (8) in order to estimate the model on the separate tax credit and subsidy samples by: (iii) BMA (with study weights); and (iv) WLS (including all observations from the omitted study – i.e. Ugur et al., 2015 – and with study weights). The corresponding results are reported in Table 2b (Section 5).

¹² The robust regression model implemented by Stata's *rreg* excludes observations for which the Cook's distance (D) is greater than unity ($D > 1$). Given that Cook's D is essentially a measure of leverage, the finding – after estimating Eq. (8) (unweighted) – that for only one observation is $D > 1$ (2.37) indicates that there is no undue problem of excess leverage. This, in turn, is reassuring for the accuracy of our coding (Stanley and Doucouliagos, 2012: 41-42 and 94). The consequent elimination of this single observation from robust regression is reflected in the number of observations recorded in Table 2a. Robust regression also up weights (down weights) observations with small (large) residuals.

¹³ BMA model estimates are derived from estimating 2^m different combinations of our m moderator variables. In these models, we restrict the essential MRA variables (see Eq.3 and Eq.4) to appear in every specification: i.e. the constant term; and the precision term. Coefficient estimates are reported as a weighted average of estimates from the 2^m models, where the weights are determined by the posterior model probabilities, which reflect the fit of the model to the data. Estimations in this study were performed by the user written Stata command, *bma* (De Luca and Magnus, 2011).

Table 1. Moderator variables with descriptions, means and standard deviations for each literature

				Study-unweighted means (standard deviation)		Study-weighted means (standard deviation)	
No.	Variables	Description	Z or K	Tax credit	Subsidy	Tax credit	Subsidy
Observations				251	323	251	323
Dependent variable							
1.	PCC (Partial Correlation Coefficient)	Partial Correlation between R&D support (tax credit or subsidy) and private R&D expenditure	n.a.	.0649 (.0624)	.0676 (.0991)	.0675 (.0581)	.0603 (.1134)
Moderator variables							
2.	Incremental-based tax-credit system	=1 when the tax-credit system is incremental-based only, 0 otherwise (i.e. if volume-based or mixed)	Z	.6335 (.4828)	n.a. ^a	.4167 (.4940)	n.a. ^a
3.	Growth	=1 when the outcome variable is measured as the growth of R&D expenditure, 0 otherwise	Z	.2709 (.4453)	n.a. ^a	.3076 (.4624)	n.a. ^a
4.	High-tech only	=1 if high or medium-high technology firms only are included in sample, 0 otherwise (i.e. low or medium-low technology firms)	Z	.2351 (.4249)	.1950 (.3969)	.2618 (.4405)	.2188 (.4140)
5.	Manufacturing sector only	=1 if firms come from manufacturing sector only, 0 otherwise	Z	.1873 (.3909)	.4768 (.5002)	.4313 (.4963)	.4722 (.5000)
6.	Starting point of data ¹⁴	=1 if the starting point of the data is in 1996 (median of the whole sample) or later, 0 otherwise	Z	.5896 (.4929)	.5356 (.4995)	.7250 (.4474)	.4618 (.4993)
7.	Panel data	=1 if study uses panel data, 0 otherwise	Z	.3546 (.4793)	.2446 (.4305)	.5903 (.4928)	.2471 (.4320)
8.	Micro & SMEs only	=1 when micro, small or medium firms only are included in the sample, 0 otherwise	Z	.0956 (.2947)	.1486 (.3563)	.0427 (.2025)	.1803 (.3850)
9.	Dynamic panel estimation	=1 when dynamic panel estimation is used, 0 otherwise	Z	.1434 (.3512)	.0402 (.1968)	.1181 (.3233)	.0938 (.2919)

¹⁴ In order to investigate the impact of the Global Financial Crisis on the effectiveness of R&D support measures, data from 2007-08 and later is required. However, only 21 such observations exist in the MRA database, which are all reported from one tax credit study.

10.	Difference-in-Differences	=1 if DiD method is used, 0 otherwise ^b	Z	.2390 (.4274)	.1548 (.3623)	.1944 (.3966)	.1117 (.3155)
11.	Instrumental variable (IV) estimation	=1 if IV estimation is used, 0 otherwise ^c	Z	.1355 (.3429)	.1146 (.3190)	.2292 (.4211)	.1993 (.4001)
12.	R&D performers only	=1 if only R&D performers are included in the sample, 0 otherwise	Z	.2749 (.4474)	.4303 (.4959)	.3667 (.4829)	.3837 (.4870)
13.	Developing economies	=1 if effect sizes refer to developing economies, 0 otherwise	Z	.1673 (.3740)	.1022 (.3033)	.3421 (.4754)	.1250 (.3312)
14.	Binary treatment variable	=1 if a binary indicator variable is used to acknowledge receipt of tax credits/subsidies, 0 otherwise (i.e. amount of support)	Z	.8008 (.4002)	.5728 (.4954)	.6597 (.4747)	.5527 (.4980)
15.	No-control of endogeneity	=1 if primary estimates come from models not controlling for endogeneity, 0 otherwise	Z	.0956 (.2947)	.0372 (.1894)	.2667 (.4431)	.0574 (.2330)
16.	Tax-dominated economies	=1 if effect sizes refer to tax-dominated economies, 0 otherwise	Z	.2590 (.4389)	n.a. ^a	.4500 (.4985)	n.a. ^a
17.	Subsidy-dominated economies	=1 if effect sizes refer to subsidy-dominated economies, 0 otherwise	Z	.6454 (.4793)	.9412 (.2357)	.4211 (.4947)	.8750 (.3312)
18.	Year of publication	=1 if studies are published in 2008 (median of the whole sample) or later, 0 otherwise	K	.5299 (.5001)	.6130 (.4878)	.6667 (.4723)	.5417 (.4990)
ELASTICITIES SUBSAMPLE (see Section 6)							
Observations				24	32	24	32
Dependent variable:							
19.	Elasticity	% response of firms' R&D spending to a 1% change in the amount of tax credit or subsidy received	n.a.	.0657 (.1128)	.0281 (.4298)	.0838 (.1166)	.0062 (.3093)
Moderator variables							
20.	Starting point of data	=1 if the starting point of the data is in 1998 (median of the elasticities subsample) or later, 0 otherwise	Z	.5000 (.5108)	.5000 (.5080)	.6666 (.4815)	.5000 (.5080)
21.	Year of publication	=1 if studies are published in 2009 (median of the elasticities subsample) or later, 0 otherwise	K	.5000 (.5108)	.6875 (.4709)	.6666 (.4815)	.6666 (.4789)

^a These moderator variables appear only in the tax credit literature.

^b The moderator variable DiD refers to both difference-in-differences with matching and difference-in-differences alone.

^c The moderator variable IV refers to both selection models using IV estimation and IV estimation alone.

5. Results and discussion

Table 2a reports the results from estimating Eq. (8) by (i) WLS (with study weights); and (ii) robust regression (without study weights) and includes the following information:

1. directly estimated moderator effects for the subsidy literature (β_{m1} in Eq.8);
2. directly estimated interaction effects (β_{m2}), which capture differential moderator effects between the two literatures (i.e. the tax credit moderator effect minus the subsidy moderator effect);
3. derived tax credit moderator effects ($\beta_{m1} + \beta_{m2}$), calculated as the paired sums of the estimated subsidy effects and the corresponding interaction effects (for an example of this calculation, see Appendix C);
4. the derived publication bias in the subsidy literature, calculated as the sum of the constant term and the coefficient on *PubYear* weighted by its study-weighted mean ($\beta_1 + \gamma_1 * \text{Study weighted mean}$) (for explanation of the study-weighted mean, see Section 4);
5. the derived differential publication bias between the two literatures ($\beta_2 + \gamma_2 * \text{Study weighted mean}$) (a statistically significant positive coefficient indicates larger publication bias in the tax credit literature than in the subsidy literature); and
6. the derived publication bias in the tax credit literature, calculated as the sum of the derived publication bias in the subsidy literature and the derived differential publication bias between the two literatures ($\beta_1 + \gamma_1 * \text{Studyweighted mean}$) + ($\beta_2 + \gamma_2 * \text{Study weighted mean}$).

Table 2b reports the results from estimating our model on the separate tax credit and subsidy samples (i.e. by omitting the tax credit dummy and corresponding interaction terms from Eq.8). Where possible (i.e. for the models estimated by WLS), we report the Ramsey test to test the null of no omitted non-linear relationships in the model.¹⁵ In each case the null is not rejected (in Table 2a, $p=.497$; and in Table 2b, $p=.206$ and $p=.312$ respectively), which supports ‘the reliability of inference’ (Spanos, 2017: 13, 16).

In Table 3, we report the main findings of our study. For both the tax credit and the subsidy literatures we derive the average publication bias and the authentic effects – beyond publication bias and controlling for heterogeneity – from each of the four models estimated. (The method of derivation is explained following Eq.6 above; see Appendix C for indicative examples of the syntax used to calculate the authentic effects and publication bias.)

To aid interpretation of the estimates reported in Tables 2a, 2b and 3, we refer to the guidelines of Doucouliagos (2011: 3), who characterises PCCs as: ‘of little (small) practical significance’ ($PCC < 0.07$); or “moderate” – of greater practical significance ($0.07 \leq PCC \leq 0.33$); or “large” ($PCC > 0.33$). On this metric, for both the tax credit and the subsidy literatures, both the study-unweighted and the study-weighted unconditional mean PCCs reported in Table 1 are ‘of little (small) practical significance’. This is also the case for the conditional or estimated authentic effects (PCCs) reported in Table 3. Moreover, most of the

¹⁵ Sometimes described as testing the null of no omitted variables but more precisely as testing the null of no omitted non-linearities in model specifications.

moderating effects identified by our multiple MRA in Tables 2a and 2b are small: disregarding statistical significance, from 64 derived tax credit and 52 directly estimated subsidy effects, respectively 56 and 41 are small while eight and 11 are moderate.

We provide a qualitative overview of our findings in Table 4, which reports findings simply as positive and statistically significant (+), negative and significant (-), or as statistically insignificant (0).¹⁶

Table 2a. MRA results – sources of heterogeneity in the tax credit and subsidy literatures

Dependent: t-statistic (<i>attt</i>)	WLS (weighted)	Robust Regression (unweighted)
	Full-sample estimation	Full-sample estimation
Subsidy Z-moderators		
<i>invsepcc</i> (inverse SE of the PCC)	-0.118*** (0.0231)	-0.138** (0.0684)
<i>invSEhigh_tech</i> (high-tech sector)	-0.0280*** (0.00522)	-0.0271 (0.0179)
<i>invSEmanufacturing</i> (manufacturing sector)	0.0393*** (0.00845)	0.0386 (0.0242)
<i>invSEt_start_1996</i> (start-point of data in 1996 or later)	0.0459*** (0.0157)	0.0440* (0.0265)
<i>invSEpanel</i> (panel data)	0.278*** (0.0150)	0.282*** (0.0469)
<i>invSEmicro_smes</i> (micro & SMEs)	-0.0186 (0.0188)	-0.0200** (0.00923)
<i>invSEshort_run</i> (short-run effects)	-0.171*** (0.0205)	-0.168*** (0.0649)
<i>invSEdid</i> (DiD method)	-0.211*** (0.0169)	-0.214*** (0.0496)
<i>invSEiv</i> (IV method)	-0.0190 (0.0147)	-0.0124 (0.0160)
<i>invSErdperformersonly</i> (R&D performers only)	-0.0240*** (0.00628)	-0.0277** (0.0117)
<i>invSEdeveloping</i> (developing economy)	-0.0194 (0.0198)	-0.0259 (0.0511)
<i>invSEbinary</i> (binary measurement of subsidy)	0.00857 (0.00884)	0.00841 (0.0111)
<i>invSEno_control_endogeneity</i> (not addressing endogeneity)	0.0343*** (0.00893)	0.0402*** (0.0137)
<i>invSEtax_domination_1</i> (tax-credit-dominated economies)	n.a.	n.a.
<i>invSEsub_domination_1</i> (subsidy-dominated economies)	0.0254 (0.0202)	0.0382 (0.0646)

¹⁶ De Luca and Magnus (2011:533) suggest the BMA counterparts: positive (+) denotes a t-statistic >1; negative (-) denotes a t-statistic <-1; and zero (0) otherwise. This is ‘related to a well-known property of R^2 (the adjusted R^2), which rises if and only if the t ratio associated with an added regressor is greater than one in absolute value’.

Z-moderators interacted with taxcredit_literature dummy (tax)		
tax_invSE (inverse SE of the PCC interacted with tax credit dummy)	0.00666 (0.0373)	0.0562 (0.0909)
tax_invSEincremental (incremental tax credit system)	0.116*** (0.0244)	0.111** (0.0512)
tax_invSEgrowth_related (outcome variable differenced)	-0.0330*** (0.00225)	-0.0325*** (0.00904)
tax_invSEhigh_tech (high-tech sector)	0.0233** (0.00987)	0.0228 (0.0181)
tax_invSEmanufacturing (manufacturing sector)	-0.0380 (0.0242)	-0.0320 (0.0271)
tax_invSEt_start_1996 (start-point of data in 1996 or later)	-0.0408** (0.0198)	-0.0531* (0.0275)
tax_invSEpanel (panel data)	-0.0919*** (0.0318)	-0.109 (0.0722)
tax_invSEmicro_smes (micro & SMEs)	0.00939 (0.0193)	0.0150 (0.0102)
tax_invSEshort_run (short-run effects)	0.160*** (0.0224)	0.157** (0.0662)
tax_invSEdid (DiD method)	0.238*** (0.0421)	0.240*** (0.0563)
tax_invSEiv (IV method)	0.0279 (0.0235)	0.0232 (0.0256)
tax_invSERdperformersonly (R&D performers only)	-0.0226*** (0.00630)	-0.0195 (0.0144)
tax_invSEdeveloping (developing economy)	-0.0223 (0.0263)	-0.00770 (0.0672)
tax_invSEbinary (binary measurement of subsidy)	0.198*** (0.0231)	0.190*** (0.0560)
tax_invSEno_control_endogeneity (not addressing endogeneity)	-0.0142 (0.00898)	-0.0189 (0.0172)
tax_invSEtax_domination_1 (tax-credit-dominated economies)	-0.0385*** (0.00534)	-0.0395 (0.0242)
tax_invSEsub_domination_1 (subsidy-dominated economies)	-0.2238*** (0.0415)	-0.243*** (0.0846)
K-moderators and constant		
yearofpublication_2008 (year of publication in 2008 or later)	2.760*** (0.326)	2.754*** (0.382)
taxcredit_literature (tax credit literature dummy)	1.847** (0.685)	1.190*** (0.386)
tax_yearofpublication_2008 (interaction of yearofpublication_2008 and taxcredit literature)	-4.095*** (0.891)	-3.843*** (0.730)
_cons (constant)	-0.973*** (0.275)	-0.791*** (0.277)
Study-effects included	yes	yes
Observations	574	573
R-squared	0.706	0.733

Ramsey test	F (3, 512) = 0.79 (p=0.4974)	n.a.
Variance Equality test (tax credit and subsidy subsamples)	F(323, 251) = 1.17 (p=0.1830)	F(322, 251) = 1.10 (p=0.4471)
Derived tax credit Z-moderators		
<i>tax_invSE+invsepcc</i> (inverse SE of the PCC)	-0.1109*** (0.0293)	-0.0815 (0.0598)
<i>tax_invSEhigh_tech+invSEhigh_tech</i> (high-tech sector)	-0.0047 (0.0084)	-0.0044 (0.0032)
<i>tax_invSEmanufacturing+invSEmanufacturing</i> (manufacturing sector)	0.0013 (0.0227)	0.0066 (0.0122)
<i>tax_invSEt_start_1996+invSEt_start_1996</i> (start-point of data in 1996 or later)	0.0051 (0.0120)	-0.0090 (0.0073)
<i>tax_invSEpanel+ invSEpanel</i> (panel data)	0.1863*** (0.0281)	0.1737*** (0.0549)
<i>tax_invSEmicro_smes+ invSEmicro_smes</i> (micro & SMEs)	-0.0092* (0.0047)	-0.0050 (0.0043)
<i>tax_invSEshort_run+invSEshort_run</i> (short-run effects)	-0.0109 (0.0092)	-0.0114 (0.0135)
<i>tax_invSEdid+ invSEdid</i> (DiD method)	0.0268 (0.0386)	0.0267 (0.0265)
<i>tax_invSEiv+ invSEiv</i> (IV method)	0.0088 (0.0184)	0.0109 (0.0200)
<i>tax_invSERdperformersonly+ invSERdperformersonly</i> (R&D performers only)	-0.0467*** (0.0005)	-0.0472*** (0.0085)
<i>tax_invSEdeveloping+invSEdeveloping</i> (developing economy)	-0.0417** (0.0172)	-0.0335 (0.0437)
<i>tax_invSEbinary+ invSEbinary</i> (binary measurement of subsidy)	0.2064*** (0.0213)	0.1989*** (0.0549)
<i>tax_invSEno_control_endogeneity+ invSEno_control_endogeneity</i> (not addressing endogeneity)	0.0201*** (0.0009)	0.0213** (0.0104)
<i>tax_invSEtax_domination_1 + invSEtax_domination_1</i> (tax-credit-dominated economies)	-0.0385*** (0.0053)	-0.0395 (0.0242)
<i>tax_invSEsub_domination_1 + invSEsub_domination_1</i> (subsidy-dominated economies)	-0.1984*** (0.0362)	-0.2049*** (0.0547)
<i>tax_yearofpublication_2008+ yearofpublication_2008</i> (Publication bias evolution)	-1.335 (0.829)	-1.089* (0.623)
Joint significance of differential MVs	F(14, 35) = 164.00 (p=0.0010)	F(14, 515) = 3.90 (p=0.0000)

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 2b. MRA results – sources of heterogeneity in the tax credit and subsidy literatures

Dependent: t-statistic (<i>attt</i>)	BMA (weighted)		WLS (weighted) - including Ugur et al. (2015)	
	Separate-sample estimation		Separate-sample estimation	
	Tax credit	Subsidy	Tax credit	Subsidy
Moderators (tax credit or subsidy)				
<i>invsepcc</i> (inverse SE of the PCC)	-0.0343 [†] (0.0319)	-0.0333 (0.0538)	0.0438 (0.0471)	-0.0305** (0.0146)
<i>invSEincremental</i> (incremental tax credit system)	0.00819 (0.0273)	n.a.	0.130*** (0.0272)	n.a.
<i>invSEgrowth_related</i> (outcome variable differenced)	-0.0158 [†] (0.0153)	n.a.	-0.0335*** (0.00263)	n.a.
<i>invSEhigh_tech</i> (high-tech sector)	-0.000225 (0.00178)	-0.00449 (0.0152)	-0.00268 (0.00948)	-0.0219*** (0.0041)
<i>invSEmanufacturing</i> (manufacturing sector)	0.000679 (0.00506)	0.00628 (0.0177)	0.000953 (0.0232)	0.0012** (0.0005)
<i>invSEt_start_1996</i> (start-point of data in 1996 or later)	0.000158 (0.00304)	0.0155 (0.0266)	-0.00935*** (0.00293)	0.0278* (0.0147)
<i>invSEpanel</i> (panel data)	0.119 [†] (0.0308)	0.200 [†] (0.0441)	0.0196 (0.0350)	0.0672*** (0.0112)
<i>invSEmicro_smes</i> (micro & SMEs)	-0.000652 (0.00336)	-0.00351 (0.00881)	-0.0106* (0.00568)	-0.0174 (0.0190)
<i>invSEshort_run</i> (short-run effects)	-0.000592 (0.00436)	-0.133 [†] (0.0508)	-0.0114 (0.00983)	-0.0181 (0.0157)
<i>invSEdid</i> (DiD method)	-0.0120 (0.0269)	-0.151 [†] (0.0502)	0.0528 (0.0498)	-0.0357*** (0.0102)
<i>invSEiv</i> (IV method)	-0.00117 (0.00778)	-0.0301 [†] (0.0239)	0.0201 (0.0204)	-0.0498*** (0.0139)
<i>invSERdperformersonly</i> (R&D performers only)	-0.0486 [†] (0.0168)	-0.00533 (0.0126)	-0.0467*** (0.000379)	-0.0220*** (0.0051)
<i>invSEdeveloping</i> (developing economy)	0.000293 (0.0181)	-0.00814 (0.0297)	-0.0338 (0.0192)	-0.1215*** (0.0210)
<i>invSEbinary</i> (binary measurement of subsidy)	0.123 [†] (0.0318)	0.00524 (0.0129)	0.0507 (0.0292)	-0.0012 (0.0015)
<i>invSEno_control_endogeneity</i> (not addressing endogeneity)	0.00603 (0.0141)	0.0250 [†] (0.0190)	0.0203*** (0.000769)	0.0010 (0.0020)
<i>invSEtax_domination_1</i> (tax-credit-dominated economies)	-0.00233 (0.00953)	n.a.	-0.0385*** (0.00535)	n.a.
<i>invSEsub_domination_1</i> (subsidy-dominated economies)	-0.0718 [†] (0.0460)	0.00345 (0.0275)	-0.230*** (0.0476)	0.1461*** (0.0101)
K-moderators and constant				
<i>yearofpublication_2008</i> (year of publication in 2008 or later)	-0.387 (0.574)	3.096 [†] (0.460)	-1.603 (0.964)	2.388*** (0.353)
<i>_cons</i> (constant)	0.408 [†] (0.352)	-1.414 [†] (0.317)	1.225 (0.781)	-0.875** (0.372)

Study-effects included	Yes	Yes	Yes	Yes
Observations	251	323	251	341
R-squared	n.a.	n.a.	0.608	0.660
Ramsey test	n.a.	n.a.	F (3,222) = 1.54 (p=0.2057)	F (3,303) = 1.20 (p=0.3117)
Number of models (BMA only)	131,072	16,384	n.a.	n.a.
Joint significance of differential MVs	n.a.	n.a.	n.a.	n.a.

† “Significant” in the BMA sense of ‘robustly correlated with the dependent variable’ (De Luca and Magnus, 2011: 533)

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3. Authentic effects (PCCs) and publication bias (PB) for the tax credit and subsidy literatures (derived from Tables 2a and 2b)

	WLS (study-weighted)	ROBUST REGRESSION (study-unweighted)	BMA (study-weighted)	WLS - including Ugur et al. (2015) (study-weighted)
	Derived from pooled-sample estimates		Derived from separate-sample estimates	
	1.	2.	3.	4.
Average tax credit effect	.067*** (.014)	.065*** (.010)	.064*** (.011)	.062*** (.016)
Average subsidy effect	.040*** (.006)	.025*** (.008)	.051*** (.012)	.039*** (.008)
(Tax credit – subsidy) effect	.027* (.015)	.040*** (.013)	n.a.	n.a.
Tax credit publication bias	-.02 (.50)	-.18 (.33)	.15 (.41)	.16 (.56)
Subsidy publication bias	.52*** (.16)	.90*** (.20)	.26 (.28)	.55** (.20)
(Tax credit – subsidy) PB	-.54 (.53)	-1.08*** (.39)	n.a.	n.a.

† “Significant” in the BMA sense of ‘robustly correlated with the dependent variable’ (De Luca and Magnus, 2011: 533)

*** p<0.01; ** p<0.05; * p<0.1

Table 4. Qualitative summary of results: sources of heterogeneity; publication bias; and authentic effects

	Tax Credit (TC)				Subsidy				Comparison (TC – Subsidy)			
	Full-sample estimation		Separate-sample estimation		Full-sample estimation		Separate-sample estimation		Full-sample estimation		Separate-sample estimation	
	WLS	RR	BMA	WLS	WLS	RR	BMA	WLS	WLS	RR	BMA	WLS
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
Observations	574	573	251	251	574	573	323	341	574	573	n.a.	n.a.
A. Moderator Variables (MVs)												
Contextual:												
Incremental-based system	+	+	0	+	n.a.				n.a.			
R&D performers only	-	-	-	-	-	-	0	-	-	0	n.a.	n.a.
Micro and SMEs only	-	0	0	-	0	-	0	0	0	0		
High-tech only	0	0	0	0	-	0	0	-	+	0		
Manufacturing sector only	0	0	0	0	+	0	0	+	0	0		
Developing economies	-	0	0	0	0	0	0	-	0	0		
Tax-dominated economies	-	0	0	-	n.a.				n.a.	n.a.		
Sub-dominated economies	-	-	-	-	0	0	0	+	-	-		
Research practices:												
R&D expenditure growth	-	-	-	-	n.a.				n.a.			
DiD	0	0	0	0	-	-	-	-	+	+	n.a.	n.a.
No control of endogeneity	+	+	0	+	+	+	+	0	0	0		
IV	0	0	0	0	0	0	-	-	0	0		
Dynamic panel estimation	0	0	0	0	-	-	-	0	+	+		
Panel data	+	+	+	0	+	+	+	+	-	0		
Binary treatment variable	+	+	+	0	0	0	0	0	+	+		
B. Publication Bias and Authentic R&D Support Effects												
Average publication bias	0	0	0	0	+	+	0	+	0	-	n.a.	n.a.
Publication bias evolution: Year of publication (MV)	0	-	0	0	+	+	+	+	-	-		
Authentic effects	+	+	+	+	+	+	+	+	+	+		
Authentic effect evolution: Start-point of data (MV)	0	0	0	-	+	+	0	+	-	-		

Key: + 've / - 've indicate positive/negative and statistically significant at the 10% level (WLS, RR); t-statistic > 1 in absolute value (BMA). 0 indicates not statistically different from zero (WLS, RR); t-statistic < 1 in absolute value (BMA).

The qualitative overview of our estimates presented in Table 4 begins by indicating sources of heterogeneity in the tax credit and subsidy literatures, respectively (Panel A). For convenience, we group these into moderator variables capturing the different contexts of primary studies and those capturing different research practices. The estimated effect of each moderator is informative about the varying effectiveness of tax credits and subsidies in these contexts or according to the research methods employed. A positive (negative) moderator effect indicates a context or research practice typically strengthening (weakening) the association between public R&D support and private R&D expenditure, making the PCC either more (less) positive or less (more) negative, other factors held constant. Where applicable, the comparison columns of Panel A report the differences between the tax credit and subsidy moderator effects, which are directly estimated by the interaction terms between the tax credit dummy and each moderator (reported in Table 2a). Direct comparison supported by significance tests is not possible for separate-sample estimation (Columns 11 and 12).

We begin our discussion with those moderator variables capturing sample heterogeneities, or contextual influences on the effectiveness of R&D support.

- According to the survey of Köhler et al. (2012), evaluations of incremental- or volume-based tax incentive schemes have not established any systematic differences in their respective effectiveness. However, we find incremental schemes to be more strongly associated with additionality than are volume or hybrid incremental/volume schemes (three from four estimates are significantly positive). This is consistent with Baghana and Mohnen (2009), who argue that incremental R&D tax credits do not suffer from deadweight loss and are thus preferable to level-based tax incentives. Similarly, OECD (2016: 112) argues that additionality may be more likely to arise from incremental than volume incentives, since the former ‘seeks to minimise the amount of “subsidised” R&D that would have been undertaken even in the absence of support’.
- Studies evaluating R&D support effects on homogeneous samples of “R&D performers only” report smaller effects: all four estimates in the tax credit literature are negative; and three from four in the subsidy literature. We illustrate the quantitative effect of sample homogeneity with the WLS estimates reported in Table 2a: other factors held constant, the tax credit PCC is reduced by .047; and the subsidy PCC by .024. A more homogeneous sample reduces differences between treated and untreated firms, which, in turn, may reduce bias stemming from (self-)selection into R&D support programmes. Hence, the smaller support effects arising from samples of “R&D performers only” are consistent with similarly smaller effects arising from econometric methods that control for the potential endogeneity of public R&D support (see below).
- Both the narrative literature review of the What Works Centre for Local Economic Growth (2015) and the meta-regression analysis of Castellacci and Lie (2015) find that tax credits are more effective in promoting R&D investment by SMEs than by larger firms. Conversely, we report two from four estimates suggesting that tax credits are less effective for micro and SMEs than for large firms (the two non-significant estimates are likewise both negative), although these effects are all “small”. We suggest the following reasons for these contrasting results. The narrative approach of the What Works Centre (2015) takes no account of publication selection bias. However, in common with the present study, Castellacci and Lie (2015) do control for publication selection bias, so in this case the contrast can be attributed to either sample differences (see Appendix A) or important methodological differences. In particular, two specification differences may account for the contrasting findings for these two moderator variables: in distinction to Castellacci and Lie (2015), the present study controls for (i) a wider range of moderator

variables and (ii) study fixed effects.¹⁷ It is likely that firm size is highly correlated with these variables, in particular the fixed effects. If so, then the effects estimated for firm size by Castellacci and Lie (2015) may be influenced by omitted variable bias, whereas the estimates in the present study are free of this potential source of omitted variable bias.

Consistent with Dimos and Pugh (2016), we find weak evidence of reduced effectiveness of subsidies for micro firms and SMEs. Finally, we find no evidence of differential effectiveness of tax credits and subsidies for micro firms and SMEs (Table 4, Columns 9 and 10).

- Whereas Castellacci and Lie (2015) find that tax credits are less effective in promoting R&D investment by high-tech firms relative to firms in lower technology categories, we find no evidence to suggest this (all four estimates are statistically insignificant). Again, this contrast is likely to be related to the methodological differences between Castellacci and Lie (2015) and the present study. Conversely, Dimos and Pugh (2016) report lower effectiveness of subsidies for high-tech firms, with which the evidence from the present study is consistent (two significantly negative estimates).
- Similar to Castellacci and Lie (2015), we find no evidence of differential effectiveness of tax credits by broad sector (manufacturing only with respect to the omitted categories, i.e. services only and manufacturing and services jointly). However, two from four estimates suggest greater additionality of subsidies to manufacturing firms.
- We find no systematic evidence of differential effectiveness according to the level of development for either tax credits or subsidies. Of the three studies (33 observations) on subsidies from developing economies one is from Turkey (Ozcelik and Taymaz, 2008) and two from eastern Germany (Alecke et al., 2012 and Almus and Czarnitzki, 2003), which these respective studies classify as developing. Accordingly, in the absence of more comprehensive evidence from developing economies, we cannot generalise from these findings.
- Finally, we estimate the effects of the two “external” moderator variables (i.e. from outside the primary literature). In comparison to countries using a balanced policy mix, we find, for the effectiveness of tax credits:
 - some evidence of reduced effectiveness of tax credits in countries using mainly tax credits (*Tax-dominated economies*) (two significantly negative estimates and one of borderline significance – $p=.103$); and
 - uniform evidence of reduced effectiveness of tax credits in countries using mainly subsidies (*Sub-dominated economies*) (four significantly negative estimates).
- For the effectiveness of subsidies:
 - in the primary literature, there are no evaluations of subsidy effects in countries using mainly tax credits to support R&D; and
 - although we find only weak evidence of increased effectiveness of subsidies in countries using mainly subsidies (one significantly positive estimate), we find that subsidies are more effective than tax credits in economies that predominantly rely on subsidies (both comparisons – in Columns 9 and 10 – are significantly negative, indicating that the tax credit effect is significantly smaller than the subsidy effect).

We now turn to research practices that influence the size of the effects reported in the primary literature.

¹⁷ See Appendix A for a comparison of the present study with recent meta-regression studies of, respectively, the tax credit and subsidy literatures: Castellacci and Lie, 2015; and Dimos and Pugh, 2016.

- Whereas Marino et al. (2016: 14) find a strong additional effect of the value of R&D tax credits or subsidies on the value of private R&D expenditure, they report the opposite – ‘notable crowding out’ – when the treatment and outcome variables are differenced. Our findings endorse this conclusion: all four estimates are significantly negative, confirming that R&D expenditure growth effects tend to be smaller than R&D expenditure (levels) effects.¹⁸ Marino et al. (2016: 14) draw an important methodological conclusion: ‘Because differences in R&D growth account better for firm specific time-invariant effects, we are more confident in evaluations having such an outcome variable.’
- Not controlling for the potential endogeneity of public support inflates reported effect sizes: three from four tax credit effects and three from four subsidy effects are significantly positive. We conjecture that while subsidies are subject to both selection by programme managers and self-selection, firms’ self-selection may be a particularly strong influence on applications for tax credits.
- Compared to other estimation approaches, difference-in-differences (DiD) and instrumental variable estimation (IV) – approaches known to be effective in controlling for both observable and unobservable sources of endogeneity – tend to reduce subsidy effects (all four DiD effects and two from four IV effects are significantly negative). This is consistent with the possibilities for selection bias noted above. However, neither approach to estimation has a discernible tax credit effect. In no case is either approach associated with a positive effect, while the predominance of insignificant results may reflect domination of the omitted category by matching approaches. (Although not controlling for unobservable influences on selection, matching does mitigate potential endogeneity by controlling for observable influences and thus sets the bar rather high.) Our findings on DiD and IV estimation are broadly consistent with those reported by Castellacci and Lie (2015) and Dimos and Pugh (2016).

We now have four types of evidence each suggesting the overriding importance of addressing the potential endogeneity of the selection of firms into R&D support: (i) homogeneous samples of R&D performers; (ii) specification of R&D expenditure in growth rather than in levels; (iii) estimation approaches that control for endogeneity; and (iv) estimation approaches that control for both observed and unobserved influences on the selection process (compared to all other approaches, including – mainly – matching). This evidence uniformly supports the consensus on the need to control for the potential endogeneity of subsidy support, and supports the current practice of also treating tax credits as potentially endogenous (Becker, 2015; Czarnitzki et al., 2011; Yang, et al., 2012).¹⁹

¹⁸ All 68 estimates of growth effects – from 6 studies – are from the tax credit literature.

¹⁹ To obtain comparable authentic representative effects from the tax credit and subsidy literatures, our meta-regression models thus address two types of selection bias: (i) by specification with moderator variables to measure and control for the *observed* presence or absence of sample, specification and estimation practices in the primary studies that address (or do not address) selection and/or self-selection of firms into R&D support; and (ii) by controlling for publication selection bias (also known as selective reporting bias), which is the aggregate bias in the effect sizes reported by an econometric literature arising from the individual *unobserved*

In addition, we report three other influences from research practices.

- There is strong evidence that the use of *panel data* has a positive effect in both literatures, which is consistent with Dimos and Pugh (2016: 808) who suggest that this effect may occur because panel data enables researchers to ‘capture cumulative effects over time’.
- *Dynamic panel estimation* captures the effect of reporting coefficients from dynamic panel model estimation. We find no effect on the tax credit estimates but we do find evidence of negative effects on the subsidy estimates (three from four estimates). One reason for the absence of positive effects and presence of negative effects arises from the interpretation of directly estimated coefficients in dynamic models as short-run or impact effects, which cannot be larger than the effects reported from static models.
- There is some discussion of the effect of using binary indicators for R&D support rather than continuous data on the value of support (Görg and Strobl, 2007). Consistent with Dimos and Pugh (2016), we find that using binary indicators does not systematically influence subsidy additionality effects. However, we find that this choice does matter for the tax credit literature, where effect sizes may be inflated by using a binary indicator rather than the value of the tax credit (three significantly positive effects; Castellacci and Lie, 2015, do not investigate this effect).²⁰ This is in line with the argument of Hall and Van Reenen (2000: 458) who identify the main disadvantage of the “Additionality Ratio” literature analysed by the present MRA: the use of dummy variables to identify tax credit effects means ‘that the measurement is relatively imprecise, because there is no guarantee that all firms are facing the same magnitude of credit at any given point in time’. The implication is that future evaluations of the effectiveness of tax credits should follow the example of studies that use the actual value of tax credits rather than a binary indicator. In our sample, 20 per cent of total estimates conform to this best practice guideline; see Table 1.

We find no evidence of publication bias (*PB*) in the tax credit literature throughout the sample period. Conversely, “little to modest” positive publication bias (i.e. $|PB| < 1$; Doucouliagos and Stanley, 2013) is found in three of the four estimates from the subsidy literature. Moreover, although we find at most weak evidence that publication bias has decreased over time in the tax credit literature (one significantly negative estimate), we find uniform evidence of increasing publication bias in the subsidy literature. This contrast between the two literatures is confirmed by both comparisons (Columns 9 and 10), which indicate that the evolution of publication bias in the subsidy literature is significantly larger than in the tax credit literature. Contrary to the tax credit literature, which throughout the sample period has not been contaminated by publication bias, the subsidy literature displays the characteristics of the well-known “decline effect”, whereby many ‘scientifically discovered effects published in the literature seem to diminish over time’ (Schooler, 2011:

specification searches, in each primary study, that inform the selection of the estimates to be submitted for publication.

²⁰ There are advantages in using a continuous over a binary measurement of public support receipt. Continuous data enable identification not only of non-linear effects in both literatures but also of different degrees of crowding out for the subsidy literature (Dimos and Pugh, 2016).

437). Early evaluations of R&D subsidies tended to report crowding-out effects, even full crowding out (Wallsten, 2000). However, according to the decline effect, initially large supportive findings over time give way to smaller and even contradictory findings. In the particular case of the subsidy literature, whereas the initial novelty was the finding of crowding out, we conjecture that, subsequently, the search for novelty increasingly favoured less negative and, eventually, positive findings (additionality).

Our main findings are the authentic effects estimated for both literatures after controlling for publication bias and heterogeneity. For both the tax credit and the subsidy literatures the authentic effects are significantly positive in all cases. Measured as PCCs, the overall authentic tax credit effects vary between .062 and .067, while the subsidy effects vary between .025 and .051 (see Table 3). The high degree of precision of these estimates (all statistically significant at the 1% level or lower) supports the validity of reporting overall, representative effects for both tax credits and subsidies, despite the many heterogeneities revealed by the moderator variables.²¹ There are two further striking features of these results: first, both comparisons reported in Table 4 (Columns 9 and 10) indicate that tax credit effects are more strongly correlated with firms' R&D expenditure than are subsidy effects (respectively by .027, $p=.086$, and .040, $p=.002$; see Table 3); yet, second, these comparisons should be interpreted in the context that all of these effects are “small”, according to the guidelines introduced above.

Finally, our findings provide partial support for Becker's (2015) conjecture that reported R&D support effects are tending to grow over time.²² Our “time effect” moderator divides studies in the primary literature according to the starting year of their datasets. Whereas tax credit effects have been of similar effectiveness throughout the sample period (three non-significant estimates), our estimated subsidy effects show a tendency to increase over time (three significantly positive estimates). Therefore, given that we take into account changes in publication selection over time, we identify evolving subsidy effects *beyond* publication bias. We conjecture that tax credit effects do not change substantially over time, because tax credits are non-discretionary, delivered at arm's length subject to rules that, while changeable from time to time, are designed to give businesses a predictable environment for their R&D decisions. Conversely, subsidies have to be targeted, which requires prior identification of investments with social returns above private returns, designed according to widely varying business contexts, and then implemented. Each stage in this (simplified) policy cycle requires decisions informed by intense engagement with potentially eligible firms, and thus a prolonged process of diverse feedback and learning. If so, greater requirement and scope for learning may account for increased subsidy effects over time.

²¹ The robustness of the estimated authentic effects for both the tax credit and subsidy literatures across different approaches to estimation and different samples (Table 3) extends also to the estimated moderator effects (Tables 2a, 2b and 4). For example, of the 35 effects estimated from both the pooled data excluding the results from Ugur et al. (2015) and the separate samples including the results from Ugur et al. (2015), 26 are both statistically significant and of the same sign, while all but one of the remaining nine have the same sign.

²² Dimos and Pugh (2016: 808) make a similar point for the subsidy literature, finding ‘positive effects on reported effect sizes of using more recent data, which is consistent with increasing effectiveness over time of subsidies’. In contrast, Castellacci and Lie (2015: 826) find that tax credit studies ‘published after the year 2000 have on average reported a lower additionality ratio’.

The next Section extends our analysis from policy effects measured as *PCCs* to policy effects measured directly in value terms (i.e. to indicate by how much the R&D expenditure changes as a result of R&D support).

6. MRA of elasticities subsample

To gain an indication of the size of tax credit and subsidy effects, we follow Dimos and Pugh (2016) in exploring subsamples from both literatures that report their findings as constant elasticities, i.e. unit-invariant measures. Elasticities – the percentage response of firms' R&D spending to a one per cent change in the amount of tax credit or subsidy received – from the primary studies yield comparable effect sizes without transformation. The combined subsample has 56 estimates, 24 from the tax credit and 32 from the subsidy literature. Descriptive statistics for the elasticities are presented in Table 1 (Row 19). Although the sample size is relatively small, our analysis benefits from both a general principle underlying MRA – namely, ‘the increase in statistical power of hypothesis testing when ... pooling study outcomes’ (Koetse et al., 2010: 218) – and our specific approach to gaining efficiency by using interaction terms rather than separate sample estimation.

MRA on the elasticity subsample can be conducted by analogy with Eqs (3) – (5) where, instead of the PCC, our dependent variable is the untransformed elasticity (e) reported in the primary literature:

$$e_i = \beta_0 + \beta_1 SE_i^e + \varepsilon_i \quad (9)$$

$$\frac{e_i}{SE_i^e} = t_i^e = \beta_1 + \beta_0 \left(\frac{1}{SE_i^e} \right) + v_i \quad (10)$$

$$t_i^e = \beta_1 + \beta_0 \left(\frac{1}{SE_i^e} \right) + \beta_2 D_i^{tax} + \beta_3 \left(\frac{1}{SE_i^e} D_i^{tax} \right) + v_i \quad (11)$$

Whereas in our previous analysis coefficients are interpreted in terms of PCCs, in this Section they are interpreted as elasticities. As in our previous analysis, we control for the evolution of publication selection to take into account Becker's (2015) conjecture that estimates of R&D support effects are rising over time. Accordingly, we augment Eq. (11) with two moderator variables to model potential time effects (Table 1, Rows 20 and 21):

1. t_start_1998 , a Z-moderator which captures the evolution of the authentic effect, and is defined as 1 for reported estimates from datasets starting in 1998 (the median) or after, and 0 otherwise; and

2. *yearofpublication_2009*, a K-moderator which captures the evolution of publication bias, and is defined as 1 for reported estimates from studies published in 2009 (the median) or later, and 0 otherwise.²³

$$\begin{aligned} \frac{e_i}{SE_i^e} = t_i^e = & \beta_1 + \beta_0 \left(\frac{1}{SE_i^e} \right) + \beta_2 D_i^{tax} + \beta_3 \left(\frac{1}{SE_i^e} D_i^{tax} \right) + \beta_4 \left(\frac{1}{SE_i^e} t_start_1998 \right) \\ & + \beta_5 \left(\frac{1}{SE_i^e} D_i^{tax} t_start_1998 \right) + \beta_6 (yearofpublication_2009) \\ & + \beta_7 (D_i^{tax} yearofpublication_2009) + v_i \end{aligned} \quad (12)$$

β_4 captures the moderating effect of time on the subsidy effect, β_5 measures the difference between the time effect in the tax credit literature and the time effect in the subsidy literature, and the sum $\beta_4 + \beta_5$ captures the moderating effect of time on the tax credit effect. The moderating effect of time on publication bias is captured by β_6 in the subsidy literature, β_7 captures the difference between the tax credit and subsidy time effects on publication bias, and the sum $\beta_6 + \beta_7$ captures the moderating effect of time on publication bias in the tax credit literature.

Eqs (10) – (12) are estimated as weighted least squares regressions. In view of the limited degrees of freedom, we do not include additional moderator variables. Whereas Eq. (11) is a restricted model, treating both publication bias and elasticities as time-invariant, Eq. (12) allows these to be time-varying and in this sense is unrestricted. Because Becker's conjectured time pattern relates to the actual effects reported in the literature, rather than to the strength of association measured by PCCs, the elasticities subsample is a particularly suitable context for this line of enquiry, despite its small size.

Table 5 presents the estimated authentic empirical effects and publication bias derived post estimation from Eq. (12). (The regression results from estimating Eq. (12) are reported in Appendix D; the method of calculating the derived effects is described in Appendix C.) The small number of studies (six from the subsidy and three from the tax credit literature) means that cluster-robust standard errors are not valid and that best practice is to report default standard errors, which is consistent with our principle of adopting a conservative approach to inference.²⁴ We also report both study-unweighted and study-weighted estimates (the first giving each estimate equal weight and the second giving every study equal weight).

²³ In both cases, taking the median across the pooled elasticities sample is a compromise between different median dates in each of the two literatures.

²⁴ For the estimates reported in Table 5, cluster-robust standard errors are typically smaller than the default standard errors reported.

Table 5. Authentic Effects and Publication Bias - constant elasticities

	Restricted Model – Eq. (11) (WLS) ^a		Unrestricted Model – Eq. (12) (WLS) ^b	
	1.	2.	3.	4.
	Study-Unweighted	Study-Weighted	Study-Unweighted	Study-Weighted
Tax credit effect				
Elasticity	0.012 (0.013) [p=0.352]	0.008 (0.015) [p=0.613]	0.017 ^c (0.011) [p=0.137]	0.012 (0.012) [p=0.301]
Tax credit effect evolution	n.a.	n.a.	-0.030 (0.023) [p=0.194]	-0.030 (0.025) [p=0.236]
Subsidy effect				
Elasticity	0.003** (0.001) [p=0.020]	0.003** (0.002) [p=0.038]	0.016*** (0.005) [p=0.001]	0.014*** (0.003) [p=0.000]
Subsidy effect evolution	n.a.	n.a.	0.018* (0.009) [p=0.051]	0.015*** (0.005) [p=0.005]
Tax credit publication bias				
Publication bias	1.10* (0.63) [p=0.087]	1.52** (0.73) [p=0.044]	0.82 (0.57) [p=0.155]	1.28** (0.56) [p=0.028]
Tax credit PB evolution	n.a.	n.a.	2.77** (1.13) [p=0.018]	2.77** (1.28) [p=0.035]
Subsidy publication bias				
Publication bias	0.73* (0.39) [p=0.068]	1.12*** (0.38) [p=0.005]	-0.07 (0.41) [p=0.863]	0.28 (0.32) [p=0.391]
Subsidy PB evolution	n.a.	n.a.	2.55** (.98) [p=0.012]	2.76*** (.62) [p=0.000]
Observations	56	56	56	56
Adjusted R ²	0.07	0.04	0.29	0.45
Ramsey test	F (3,49)=2.90 (p=0.044)	F(3,49)=5.18 (p=0.003)	F(3,45)=0.41 (p=0.745)	F(3,45)=0.95 (p=0.424)

^a Columns 1 & 2 – elasticities and publication bias directly estimated; ^b Columns 3 & 4 – elasticities and publication bias derived (see Appendix C)

Default standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; ^c Statistically significant at the 10% level (one-tail test).

Table 5 reports results not only from our preferred model (Eq.12) but also from the restricted model (Eq.11) in order to highlight the benefits of controlling for potential time variation in the authentic empirical effects and publication bias: i.e., benefits with respect to (i) statistical validity; (ii) explanatory power and corresponding precision of the estimated effects; and (iii) more informative economic interpretation.

According to the Ramsey test, our unrestricted models – both study-unweighted and study-weighted – are more satisfactory with respect to their statistical properties ($p = 0.745$ and 0.424 , respectively) than are the restricted models, both of which reveal unmodelled non-linearities ($p=0.044$ and $p=0.003$, respectively).

In our preferred models (Columns 3 and 4), we do not restrict either publication bias or the elasticities estimated beyond publication bias to be constant over time. In these unrestricted models, in spite of the loss of degrees of freedom in the context of a small sample, we find that the estimated elasticities are both larger and more precisely estimated than those arising from the corresponding restricted models. In the unrestricted models, the estimated elasticities of firms' R&D spending with respect to R&D subsidy in both the study-unweighted (Column 3) and in the study-weighted models (Column 4) are much larger than in the respective restricted models (Columns 1 and 2) and statistically significant at the one per cent level.²⁵ In comparison, the estimated tax credit elasticities from the unrestricted models are also larger, although the differences between the unrestricted and the restricted models for the tax credit literature are much smaller than for the subsidy literature. The gain in precision is consistent with the increase in explanatory power – in the study-weighted estimates the Adjusted R^2 increases to .45 compared to .04 in the restricted model.²⁶

The preferred or unrestricted model estimates are more comparable across the two literatures than are the restricted model estimates: namely, elasticities of .017 (Column 3) and .012 (Column 4) for the tax credit literature; and .016 and .014 for the subsidy literature. Both subsidy estimates are statistically significant at the one per cent level. In the tax credit literature, less precise estimation reflects small sample size (24 observations). However, if we exclude the possibility of a negative tax credit effect on theoretical grounds, then a one-tail test is appropriate (as we argue in Section 2 above). In this case, the estimated positive tax credit elasticity becomes statistically significant at the 10% level in the study-unweighted model although not in the study-weighted model.

Both the tax credit and the subsidy literatures yield authentic elasticities similar in size (around .015 in all cases). Two of the three tax credit studies provided sufficiently detailed descriptive statistics to calculate an average R&D tax credits to private R&D expenditure ratio of 0.23, which is consistent with international evidence presented in OECD (2015: 171), while four from the six subsidy studies provide sufficient information to calculate an average R&D subsidy to private R&D expenditure ratio of 0.19. Hence, across the two literatures, we

²⁵ Dimos and Pugh (2016) estimate the study-unweighted restricted model and report a subsidy elasticity of .003, which we replicate in Table 5, Column 1), whereas our study-unweighted unrestricted model with time effects yields an estimate of .016. Although both elasticities are small, the earlier estimate can be interpreted as 'economically negligible' (Dimos and Pugh, 2016:811) whereas our estimate can be characterised as small but non-negligible.

²⁶ The standard errors on regression coefficients are inversely related to the goodness of fit of the regression from which they are estimated.

assume a round terms representative support to expenditure ratio of 0.2. The elasticity of public support (e) is the percentage response of private R&D expenditure to a percentage change in R&D support – hence the product of the ratio of the change in expenditure to the change in support and the ratio of support to expenditure.²⁷ Hence, an indicative elasticity of .015 and an indicative support to expenditure ratio of .200 together imply an additional private expenditure of \$.075 for every additional \$1 of public support received (.015 × [1/.2] = .075). This is very close to Yang et al. (2012: 1586) who find that ‘a one-dollar taxation remit of R&D tax credit induces .094 dollars more of R&D expenditure’.

Allowing for time-varying effects not only has a control function but also yields additional information. For the tax credit literature, we cannot make confident judgements about the evolution of estimated elasticities beyond publication bias, because the derived time effect is statistically insignificant. In the subsidy literature, the directly estimated time effect indicates that the estimated elasticities display a statistically significant increase over time even after accounting for evolving publication bias, which is consistent with Becker’s conjecture.

7. Conclusion

This Meta-Regression Analysis (MRA) compares the effectiveness of R&D tax credits and R&D subsidies in promoting private R&D investment. Comparative MRA yields statistically significant estimates of the authentic PCC between public R&D support and private R&D expenditure varying between .062 and .067 for the tax credit literature and between .025 and .051 for the subsidy literature. Although the tax credit effects are significantly more strongly correlated with firms’ R&D expenditure than are subsidy effects, these effects are all “small” according to standard guidelines. While the PCCs are “small”, supplementary MRA of the studies that report their findings as constant elasticities indicates effects that are nonetheless economically non-negligible: in round terms, an additional \$1 of public R&D support of either type induces 7.5 cents of additional private R&D expenditure. These supplementary findings confirm that both measures yield input additionality and indicate that their effects are quantitatively similar. For comparison, Foreman-Peck (2013: 64) in his study of UK SMEs finds no significant difference between the effects of R&D tax credits and other ‘non-tax-credit innovation aid’ on the propensity to innovate (a measure of output additionality).

In MRAs of both our PCC sample and our elasticities subsample, a feature of our strategy to identify these R&D support effects is to control for time-variation both in publication selection bias and in the authentic effects. We control not only for publication bias, which is typical practice in MRA, but also for its potential evolution, thereby increasing confidence in the validity of our estimated “authentic empirical effects” not only “beyond publication bias” (Stanley, 2005) but also “beyond” potential time-variation in publication bias. In addition, controlling for time-variation in the “authentic empirical effects” yields evidence on Becker’s (2015) conjecture that the reported effects of R&D support are tending to increase over time. For the tax credit literature, we find no such evolution; whereas we find evidence that subsidy effects in terms of both PCCs and elasticities are increasing over time. We conjecture that

²⁷ $e = \frac{\% \Delta(R\&D \text{ expenditure})}{\% \Delta(R\&D \text{ support})} = \frac{\Delta(R\&D \text{ expenditure})}{\Delta(R\&D \text{ support})} \times \frac{(R\&D \text{ support})}{(R\&D \text{ expenditure})}$

while tax credits are non-discretionary and delivered at arm's length, subsidies entail intense engagement with applicant firms, which may give more scope for learning.

Our findings provide evidence on heterogeneous effects in different contexts that may inform the choice between tax credits or subsidies. Although neither instrument systematically outperforms the other, analysis of moderating influences shows that in different contexts (i) the effectiveness of both tax credits and subsidies varies and (ii) there is differential effectiveness between the two instruments. Tax credits are less effective for micro firms and SMEs than for large firms, and may be more effective in economies with a balanced “policy-mix” regime rather than with either tax credits or subsidies as the dominant approach. Moreover, “incremental” schemes are the most effective way to deliver tax credits, consistent with their more stringent eligibility criteria. Subsidies are more effective than tax credits in economies predominantly using subsidies, although we find only weak evidence of increased subsidy additionality in such economies; and, generally, may be more effective for manufacturing firms, although not for high-tech firms. However, subsidies may take more time than tax credits to realise their potential, as public agencies and programme managers learn to implement them through interaction with beneficiary firms.

The choice between the two instruments does not have to be based merely upon their relative overall effectiveness or even upon their relative context-specific effectiveness. Policy makers need also to consider the returns from the induced R&D, because the two policy instruments potentially induce R&D of different quality and characteristics and, hence, with different private and social returns. While tax credits tend to promote applied R&D, subsidies can be used to target basic R&D capable of generating knowledge spillovers (Bloom et al., 2019). Accordingly, \$1 of additional R&D induced by subsidy might yield greater social returns than \$1 of additional R&D induced by tax credits. In conclusion, although we find the effectiveness of the two instruments to be similar, we caution against drawing the direct implication that tax credits and subsidies are perfect substitutes regardless of the nature of the supported R&D.²⁸

Our findings also have implications for future evaluations of R&D support. Besides a variety of strategies to address the potential endogeneity of public support measures, and using the actual value of support rather than a binary indicator (especially in tax credit studies), our results suggest the need for partnership between researchers and public authorities. To enhance the statistical power of future evaluations (especially of subsidies), small effect sizes imply the need for large datasets, which could be satisfied by access to administrative data. Analysis and discussion on the statistical power of the two literatures is presented in Appendix E.

A limitation of most MRAs is that there are sources of heterogeneity in most empirical literatures, especially those introduced by emergent themes, ‘that are associated with too few studies to be investigated’ (Dimos and Pugh, 2016: 808). In this particular MRA, we have accordingly been unable to discuss non-linear effects (Görg and Strobl, 2007; Aschhoff, 2009), policy mix – i.e. joint tax credit and subsidy – effects (Bérubé and Mohnen, 2009), the conjecture of ‘a different time pattern of the effects of tax credits and direct subsidies’

²⁸ The same observation applies also to studies in the primary literature: researchers evaluate the effectiveness of R&D tax credits and R&D subsidies in terms of their quantitative impact on firms’ R&D but typically do not focus on the returns from the supported R&D.

(Becker, 2013: 27), firm life-cycle effects (Chiang et al., 2012), and the impact of the Global Financial Crisis (Hud and Hussinger, 2015).

Acknowledgements

We received important suggestions from participants at the September 2016 Colloquium of the Meta-Analysis of Economics Research Network (MAER-Net) at Hendrix College, Arkansas, USA. Particular thanks go to Dr Dragana Radicic and Professor Mehmet Uğur who commented on the pre-submission draft. Remaining shortcomings are the authors' responsibility. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Alecke, B., Mitze, T., Reinkowski, J., Untiedt, G., 2012.** Does firm size make a difference? Analysing the effectiveness of R&D subsidies in East Germany. *German Economic Review* 13(2), Verein für Socialpolitik, 174–195.
- Almus, M., Czarnitzki, D., 2003.** The effects of public R&D subsidies on firms' innovation activities: the case of Eastern Germany. *Journal of Business and Economic Statistics* 21(2), 226–236.
- Arrow, K.J., 1962.** Economic welfare and the allocation of resources for invention. In: Nelson, R.R. (Ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton University Press, Princeton, 609–625.
- Aschhoff, B., 2009.** The effect of subsidies on R&D investment and success: do subsidy history and size matter?, ZEW Discussion Papers 09-032, ZEW—Zentrum für Europäische Wirtschaftsforschung/Center for European Economic Research.
- Becker, B., 2013.** The Determinants of R&D Investment: A Survey of the Empirical Research. Loughborough University School of Business and Economics, Working Paper 2013-09.
- Becker, B., 2015.** Public R&D Policies and Private R&D Investment: A Survey of the Empirical Evidence. *Journal of Economic Surveys*, 29(5), 917-942.
- Bérubé, C., Mohnen, P., 2009.** Are firms that receive R&D subsidies more innovative? *Canadian Journal of Economics* 42(1), 206-225.
- Baghana, R., Mohnen, P., 2009.** Effectiveness of R&D tax incentives in small and large enterprises in Québec. *Small Business Economics* 33(1), 91-107.
- Bloom, N., Van Reenen, J., Williams, H. (2019).** A Toolkit of Policies to Promote Innovation. *Journal of Economic Perspectives* 33(3) (Summer), 163-84.
- Brambor, T., Clark, W., Golder, M., 2006.** Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, 14, 63-82.

- Brodeur, A., Lé, M., Sangnier, M., Zylberberg, Y., 2016.** Star Wars: The Empirics Strike Back. *American Economic Journal: Applied Economics* 8(1), 1-32.
- Busom, I., Corchuelo, B., Martínez-Ros, E., 2014.** Tax incentives... or subsidies for business R&D? *Small Business Economics* 43(3), 571-596.
- Castellacci, F., Lie, C. M., 2015.** Do the effects of R&D tax credits vary across industries? A meta-regression analysis. *Research Policy* 44(4), 819-832.
- Chiang, S., Lee, P., Anandarajan, A., 2012.** The effect of R&D tax credit on innovation: A life cycle analysis. *Innovation: Management, policy and practice* 14(4), 510-523.
- Cohen, J., 1965.** Some statistical issues in psychological research. In: *Handbook of Clinical Psychology* (B.B. Wolman, ed.), 95–121. New York: McGraw-Hill.
- Czarnitzki, D., Hanel, P., Rosa, J.M., 2011.** Evaluating the impact of R&D tax credits on innovation: A microeconomic study on Canadian firms. *Research Policy* 40(2), 217-229.
- David, P., Hall, B., Toole, A., 2000.** Is Public R&D a Complement or Substitute for Private R&D? A review of the Econometric Evidence. *Research Policy* 29(4-5), 497-529.
- Dechezleprêtre, A., Einiö, E., Martin, R., Nguyen, K., and Van Reenen, J., 2016.** Do tax incentives for research increase firm innovation? An RD design for R&D, CEP Discussion Paper No 1413. URL (accessed 05-08-2018): <http://cep.lse.ac.uk/pubs/download/dp1413.pdf>
- De Luca, G., Magnus, R., 2011.** Bayesian model averaging and weighted-average least squares: Equivariance, stability, and numerical issues. *The Stata Journal* 11(4), 518–544.
- Dimos, C., Pugh, G., 2016.** The effectiveness of R&D subsidies: A meta-regression analysis of the evaluation literature. *Research Policy* 45(4), 797-815.
- Duguet, E., 2012.** The effect of the incremental R&D tax credit on the private funding of R&D an econometric evaluation on French firm level data. *Revue d'économie politique* (122), 405-435.
- Doucoulagos, H., 2011.** How Large is Large? Preliminary and Relative Guidelines for Interpreting Partial Correlations in Economics. School Working Paper of Economics Series, Deakin University, Melbourne.
- Doucoulagos, H., Stanley, T.D., 2009.** Publication selection bias in minimum-wage research? A meta-regression analysis. *British Journal of International Relations* 47 (2), 406–428.
- Doucoulagos, H., Stanley, T.D., 2013.** Are all economic facts greatly exaggerated? Theory competition and selectivity. *Journal of Economic Surveys* 27 (2), 316–339.
- Fisher, R., 1954.** *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.
- Foreman-Peck, J., 2013.** Effectiveness and efficiency of SME innovation policy. *Small Business Economics* 41(1), 55-70.
- Hall, B., Van Reenen, J., 2000.** How effective are fiscal incentives for R&D? A review of the evidence. *Research Policy* 29 (4/5), 449-469.

Ho, Y., 2006. Evaluating the effectiveness of state R&D tax credits. Doctoral Dissertation In: Graduate School of Public and International Affairs. University of Pittsburgh.

Hud, M., Hussinger, K., 2015. The impact of R&D subsidies during the crisis. *Research Policy* 44(10), 1844-1855.

Goolsbee, A., 1998. Does Government R&D Policy Mainly Benefit Scientists and Engineers? *American Economic Review* 88(2), 298-302.

Görg, H., Strobl, E., 2007. The effect of R&D subsidies on private R&D. *Economica* 74 (294), 215-234.

Greene, W., 1993 (2nd Ed.). *Econometric Analysis*. Prentice-Hall, New Jersey.

Ioannidis, J., Stanley, T., Doucouliagos, H., 2017. The Power of Bias in Economics Research. *The Economic Journal*, 127 (October), F236-F265. Doi: 10.1111/ecoj.12461.

Iršová, Z., Havránek, T., 2013. Determinants of Horizontal Spillovers from FDI: Evidence from a Large Meta-Analysis. *World Development* 42, 1-15.

Klette, T., Møen, J., Griliches, Z., 2000. Do subsidies to commercial R&D reduce market failures? Microeconomic evaluation studies. *Research Policy* 29(4-5), 471-495.

Koetse, M., Florax, R., de Groot, H., 2010. Consequences of effect size heterogeneity for meta-analysis. *Statistical Methods and Applications* 19(2), 217-236.

Köhler, C., Laredo, P., Rammer, C., 2012. The Impact and Effectiveness of Fiscal Incentives for R&D. Nesta Working Paper No. 12/01. URL (accessed 27-04-2017): https://www.nesta.org.uk/sites/default/files/the_impact_and_effectiveness_of_fiscal_incentives.pdf

Lee, C.Y., 2011. The differential effects of public R&D support on firm R&D: Theory and evidence from multi - country data. *Technovation* 31(5-6), 256 - 269.

Marino, M., Lhuillery, S., Parotta, P., 2016. Additionality or Crowding-out? An overall evaluation of public R&D subsidy on private R&D expenditure. *Research Policy* 45(9), 1715-1730.

Mohnen, P., n.d. R&D Tax Incentives. Innovation for Growth, Policy Brief No.25, European Commission. URL (accessed 27-04-2017): http://ec.europa.eu/competition/state_aid/legislation/workshop_rdi_pm_en.pdf

Necker, S., 2014. Scientific misbehaviour in economics. *Research Policy* 43(10), 1747-1759.

Nelson, R.R., 1959. The simple economics of basic scientific research. *Journal of Political Economy* 67(3), 297-306.

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

OECD, 2015. *OECD Science, Technology and Industry Scoreboard 2015: Innovation for growth and society*. OECD Publishing, Paris. URL (accessed 27-04-2017): http://dx.doi.org/10.1787/sti_scoreboard-2015-en

- OECD, 2016.** OECD Business and Finance Outlook 2016. OECD Publishing, Paris.
<http://dx.doi.org/10.1787/9789264257573-en>
- OECD, 2018. Measuring Tax Support for R&D and Innovation.** On-line portal to country reports (accessed 24-09-2018): <http://www.oecd.org/sti/rd-tax-stats.htm>
- Ozcelik, E., Taymaz, E., 2008.** R&D support programs in developing countries: the Turkish experience. *Research Policy* 37(2), 258–275.
- Peneder, M., 2008.** The problem of private under-investment in innovation: A policy mind map. *Technovation* 28(8), 518-530.
- Rosenberger, R.S., Loomis, J.B., 2000.** Panel stratification in meta-analysis of economic studies: an investigation of its effects in the recreation valuation literature. *Journal of Agricultural and Applied Economics* 32(3), 459-470.
- Schooler, J., 2011.** Unpublished results hide the decline effect. *Nature* 470, 437.
- Schumpeter, J.A., 1942.** *Capitalism, Socialism and Democracy*. Routledge, London, 82–83.
- Solow, R. M. 1956.** A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics* 70(1), 65–94.
- Spanos, A., 2017.** Mis-Specification Testing in Retrospect. *Journal of Economic Surveys*. On-line, prepublication: doi: 10.1111/joes.12200.
- Stanley, T.D., 2005.** Beyond publication bias. *Journal of Economic Surveys* 19(3), 309–345.
- Stanley, T.D., 2008.** Meta-regression methods for detecting and estimating empirical effects in the presence of publication bias. *Oxford Bulletin of Economics and Statistics* 70(1), 103–127.
- Stanley, T.D., Doucouliagos, H., 2012.** *Meta-regression analysis in economics and business*. Routledge, Oxford (July 2012).
- Stanley, T.D., Doucouliagos, H., 2013.** *Neither Fixed nor Random: Weighted Least Squares Meta-Analysis*. Economics Series 2013_1, Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.
- Stanley, T.D., Doucouliagos, H., Giles, M., Heckemeyer, J.H., Johnston, R., Laroche, P., Nelson, J., Paldam, M., Poot, J., Pugh, G., Rosenberger, R., Rost, K., 2013.** Meta-analysis of economics research: reporting guidelines. *Journal of Economic Surveys* 27(2), 390–394.
- Stiglitz, J., Greenwald, B., 2015.** *Creating a Learning Society*. Columbia University Press, New York.
- Usher, D., 1964.** The welfare economics of invention. *Economica* 31(123), 279–287.
- Wallsten, S.J., 2000.** The effects of government-industry R&D programs on private R&D: the case of the small business innovation research program. *RAND Journal of Economics* 31 (1), 82–100.

What Works Centre for Local Economic Growth, 2015. Innovation: R&D Tax Credits. Evidence Review 9, London: London School of Economics.
http://www.whatworksgrowth.org/public/files/Policy_Reviews/15-10-20-Innovation-Tax-Credits-Report.pdf

Yang, C.-H., Huang, C.-H., Hou, T. C.-T., 2012. Tax incentives and R&D activity: Firm-level evidence from Taiwan. *Research Policy* 41(9), 1578-1588.

Appendices (to be made available online)

Online Appendix A. Comparison between MRA studies

	Castellacci and Lie (2015)	Dimos and Pugh (2016)	Present Study (2021)
Scope	Single instrument study	Single instrument study	Comparison
Aim	To evaluate the effectiveness of <i>tax credits</i> in promoting private R&D	To evaluate the effectiveness of <i>subsidies</i> in promoting R&D expenditure and R&D outputs	To evaluate the <i>relative effectiveness</i> of tax credits and subsidies in promoting private R&D
Unit of analysis	Firm-level studies	Firm-level studies	Firm-level studies
Primary literature: heterogeneities	Segmentation of primary literature into two parts: (i) Additionality-ratio studies; and (ii) User-Cost elasticity studies	Segmentation of primary literature into three parts: (i) Private R&D expenditure; (ii) Total R&D expenditure; and (iii) Non-expenditure R&D outputs	Segmentation of the two primary literatures into two comparable parts: (i) Additionality-ratio Tax Credit studies; and (ii) Private R&D expenditure Subsidy studies
Number of studies	34 Tax Credit Studies	52 Subsidy studies	37 studies: Tax Credit – 12; Subsidy - 25
Number of observations	352 in two separate MRA databases used to estimate separate MRA models: (i) 221 Additionality Ratio (ii) 131 User-Cost elasticity	846 in three separate MRA databases used to estimate separate MRA models: (i) 289 Private R&D expenditure (ii) 357 Total R&D expenditure (iii) 200 Non-expenditure R&D outputs	598 in a combined dataset used to estimate joint models with interaction effects: 251 tax credit effects; and 347 subsidy effects

Number of moderators	12 (+ 2 interaction terms)	25	17 (15 for both Subsidy and Tax Credit observations + two specific variables for the tax credit observations)
Meta-Regression Model weighted by precision?	Yes	Yes	Yes
Study fixed-effects	No	Yes (but not interacted with precision)	Yes (interacted with precision)
Publication bias	Statistically significant and positive	Statistically significant and positive	Statistically significant and positive
Average genuine effect	<p>Statistically significant and positive: 0.0008-0.0009, but meaning unclear.</p> <p>Comment: The ‘true empirical effect’ (p.826) is measured – incorrectly – by the estimated coefficient on precision (INVSE). The representative effect for the whole literature must take into account not only the estimated coefficient on precision but also the moderator variables with which the precision measure is interacted (see: Stanley & Doucouliagos, 2012: 98-99).</p>	<p>Statistically significant and positive: in terms of the Partial Correlation Coefficient, 0.04 (significant at the 1% level)</p> <p>Comment: Fixed effects estimation (but controlling for study-specific influences on publication).</p>	<p>Statistically significant and positive: in terms of the Partial Correlation Coefficient, 0.04 (significant at the 1% level)</p> <p>Comment: Fixed effects estimation (but controlling for study-specific influences on authentic empirical effect).</p> <p>In addition, MRA of the subset of studies reporting elasticities suggests that an additional \$1 of public R&D support of either type induces 7.5 cents of additional private R&D expenditure.</p>
Time-varying publication bias and representative effects	No	Partially	Yes

Online Appendix B. Studies in the MRA database

Tax credit studies

Chiang, S., Lee, P., Anandarajan, A., 2012. The effect of R&D tax credit on innovation: A life cycle analysis. *Innovation: Management, policy and practice* 14(4), 510-523.

Duguet, E., 2012. The effect of the incremental R&D tax credit on the private funding of R&D: an econometric evaluation on French firm level data. *Revue d'économie politique* 2012/3 122, 405-435.

Hægeland, T., Møen, J., 2007. Input additionality in the Norwegian R&D tax credit scheme. *Statistics Norway Reports*.

Ho, Y., 2006. Evaluating the effectiveness of state R&D tax credits. Doctoral Dissertation In: Graduate School of Public and International Affairs. University of Pittsburgh.

Huang, C.-H., 2009. Three essays on the innovation behaviour of Taiwan's manufacturing firms. In: Graduate Institute of Industrial Economics. National Central University, Taiwan.

Kasahara, H., Shimotsu, K., Suzuki, M., 2014. Does an R&D tax credit affect R&D expenditure? The Japanese R&D tax credit reform in 2003. *Journal of the Japanese and International Economies* 31, 72-97.

Klassen, K.J., Pittman, J.A., Reed, M.P., Fortin, S., 2004. A cross-national comparison of R&D expenditure decisions: tax incentives and financial constraints. *Contemporary Accounting Research* 21 (3), 639–680.

Kobayashi, Y., 2014. Effect of R&D tax credits for SMEs in Japan: a microeconomic analysis focused on liquidity constraints. *Small Business Economics* 42 (2), 311-327.

Lee, C.-Y., 2011. The differential effects of public R&D support on firm R&D: Theory and evidence from multi-country data. *Technovation* 31 (5-6), 256–269.

Mercer-Blackman, V., 2008. The impact of research and development tax incentives on Colombia's manufacturing sector: what difference do they make? *IMF Working Paper* 08/178.

Paff, L.A., 2005. State-level R&D tax credits: a firm-level analysis. *Topics in Economic Analysis and Policy* 5 (1), 1-27.

Yang, C.-H., Huang, C.-H., Hou, T. C.-T., 2012. Tax incentives and R&D activity: Firm-level evidence from Taiwan. *Research Policy* 41 (9), 1578–1588.

Subsidy studies

Aerts, K., Thorwarth, S., 2008. Additionality effects of public R&D funding: ‘R’ versus ‘D’. Open Access publications from Katholieke Universiteit Leuven, Faculty of Business and Economics. Accessed 12-08-2019 from:

https://www.academia.edu/4595239/Additionality_effects_of_public_R_and_D_funding_R_versus_D.

Alecke, B., Mitze, T., Reinkowski, J., Untiedt, G., 2012. Does firm size make a difference? Analysing the effectiveness of R&D subsidies in East Germany. *German Econ. Rev.* 13(2), Verein für Socialpolitik, 174–195.

Ali-Yrkkö, J., 2005. Impact of Public R&D Financing on Private R&D. Does Financial Constraint Matter? ENEPRI Working Paper No. 30/February 2005.

Almus, M., Czarnitzki, D., 2003. The effects of public R&D subsidies on firms’ innovation activities: the case of Eastern Germany. *J. Bus. Econ. Stat.* 21 (2), 226–236.

Aschhoff, B., 2009. The effect of subsidies on R&D investment and success: do subsidy history and size matter?, ZEW Discussion Papers 09-032, ZEW—Zentrum für Europäische Wirtschaftsforschung/Center for European Economic Research.

Bloch C., Graversen E.K., 2012. Additionality of public R&D funding for business R&D—a dynamic panel data analysis. *World Review of Science, Technology and Sustainable Development* 9(2/3/4), 204–220.

Carboni, O., 2011. R&D subsidies and private R&D expenditures: evidence from Italian manufacturing data. *Int. Rev. Appl. Econ.* 25 (4), 419–439.

Clausen, T. H., 2009. Do subsidies have positive impacts on R&D and innovation activities at the firm level? *Struct. Change Econ. Dynam.* 20 (4), 239–253.

Czarnitzki, D., Fier, A., 2001. Do R&D subsidies matter? Evidence from the German service sector. ZEW Discussion Paper No.019, Mannheim.

Czarnitzki, D., Hussinger, K., 2004. The link between R&D subsidies, R&D spending and technological performance. ZEW Discussion Paper No. 056, Mannheim.

Duguet, E., 2004. Are R&D subsidies a substitute or a complement to privately funded R&D? Evidence from France using propensity score methods for non-experimental data. *Revue d’Economie Politique* 114 (2), 263–292.

Dumont, M., 2013. The impact of subsidies and fiscal incentives on corporate R&D expenditures in Belgium (2001–2009). Working Paper 01-13, Federal Planning Bureau, Belgium.

González, X., Pazó, C., 2008. Do public subsidies stimulate private R&D spending? *Res. Policy* 37(3), 371–389.

Görg, H., Strobl, E., 2007. The effect of R&D subsidies on private R&D. *Economica* 74 (294), 215–234.

- Herrera, L., Bravo Ibarra, E.R., 2010.** Distribution and effect of R&D subsidies: a comparative analysis according to firm size. *Intangible Capital* 6 (2), 272–299.
- Hottenrott, H., Lopes-Bento, C., Veugeler, R., 2017.** Direct and cross scheme effects in a research and development subsidy program. *Res. Policy* 46(6), 1118–1132.
- Hussinger, K., 2008.** R&D and subsidies at the firm level: an application of parametric and semiparametric two-step selection models. *J. Appl. Econometrics* 23 (6), 729–747.
- Kaiser, U., 2006.** Private R&D and Public R&D subsidies: microeconomic evidence from Denmark. *Nationaløkonomisk Tidsskrift/Danish Journal of Economics* 144(1), 1–17.
- Klette, T. J., Moen, J., 2012.** R&D Investment Responses to R&D Subsidies: A Theoretical Analysis and a Microeconomic Study. *World Review of Science, Technology and Sustainable Development* 9(2/3/4), 169–203.
- Koga, T., 2005.** R&D subsidy and self-financed R&D: the case of Japanese high-technology start-ups. *Small Bus. Econ.* 24 (1), 53–62.
- Lach, S., 2002.** Do R&D Subsidies Stimulate or Displace Private R&D? Evidence from Israel. *J. Ind. Econ.* 50 (4), 369–390.
- Ozcelik, E., Taymaz, E., 2008.** R&D support programs in developing countries: the Turkish experience. *Res. Policy* 37, 258–275.
- Suetens, S., 2002.** R&D subsidies and production effects of R&D personnel Evidence from the Flemish Region. CESIT (Centre for the Economic Study of Innovation and Technology) Discussion paper No 2002/03, November 2002.
- Ugur, M., Trushin, E., Solomon, E., 2015.** UK and EU subsidies and private R&D investment: Is there input additionality? MPRA Paper No. 68009. <https://mpra.ub.uni-muenchen.de/68009/> (accessed 01-10-2018)
- Wallsten, S.J., 2000.** The effects of government-industry R&D programs on private R&D: the case of the small business innovation research program. *RAND J. Econ.* 31 (1), 82–100.

Online Appendix C: Example of the calculation of derived tax credit effects

The examples below reproduce syntax from our Stata Do files.

Derived tax credit moderator effects (see Table 2a):

To calculate the derived precision effect (and its standard error) for the tax credit literature ($tax_invSE + invsepcc$) in the first WLS model (Column 1) we use Stata's post-estimation *lincom* command to obtain the linear sum of the subsidy precision effect ($invsepcc$) and the corresponding tax credit interaction term (tax_invSE); i.e. $-0.118*** + 0.00666 = -.1109***$ (where *** denotes significance at the 1% level). The derived effects of the Z moderators are calculated in the same way: for example, the micro and SME tax credit effect is given by the corresponding subsidy effect ($invSEmicro_smes$) plus the corresponding tax credit interaction term ($tax_invSEmicro_smes$); i.e. $-0.0186 + 0.00939 = -.0092*$ (where * denotes significance at the 10% level).

Method of calculating publication bias for the tax credit and subsidy literatures (see Table 3):

For the subsidy literature, publication bias is derived by adding the constant term and the coefficient on the “*yearofpublication_2008*” moderator variable weighted by its study-weighted mean (*swm*) in the subsidy literature (i.e. $cons + yearofpublication_2008 * swm_{yearofpublication_2008}^{sub}$), while the tax credit PB is derived from the following sum:

$cons + taxcredit_literature + yearofpublication_2008 * swm_{yearofpublication_2008}^{tax} + tax_yearofpublication_on_2008 * swm_{yearofpublication_2008}^{tax}$, where $swm_{yearofpublication_2008}^{tax}$ is the study-weighted mean of tax credit estimates published in 2008 or later. Variable names are those that appear in Tables 2a and 2b.

Method of calculating the authentic effects for the tax credit and subsidy literatures (see Table 3):

To calculate the authentic subsidy effect together with its standard error (Table 3, Column 1) for the first WLS model (Table 2a, Column 1), we sum the estimated coefficient on the precision effect in the subsidy literature ($invsepcc$) and the coefficient on each Z moderator variable weighted by its study-weighted mean: i.e. $lincom invsepcc + invSEhigh_tech*.21875 + invSEmanufacturing*.472222 + invSEt_start_1996*.4618056 + invSEpanel*.2471264 + invSEmicro_smes*.18029 + invSEshort_run*.09375 + invSEdid*.1117098 + invSEiv*.1993056 + invSERdperformersonly*.3836806 + invSEdeveloping*.125 + invSEbinary*.5526961 + invSEno_control_endogeneity*.0574371 + invSEtax_domination_1*0 + invSEsub_domination_1*.875 + invSEstudy_1*.0416667 + invSEstudy_2*.0416667 + invSEstudy_3*.0416667 + invSEstudy_4*.0416667 + invSEstudy_5*.0416667 + invSEstudy_6*.0416667 + invSEstudy_7*.0416667 + invSEstudy_9*.0416667 + invSEstudy_10*.0416667 + invSEstudy_11*.0416667 + invSEstudy_12*.0416667 + invSEstudy_14*.0416667 + invSEstudy_16*.0416667 + invSEstudy_18*.0416667 + invSEstudy_19*.0416667 + invSEstudy_20*.0416667 + invSEstudy_23*.0416667 + invSEstudy_36*.0416667$. Note that the first WLS model is study weighted, giving each study equal weight in the regression. Hence, the weights on the subsidy study effects – $invSEstudy_1 - invSEstudy_36$ are the same.

This sums to the average subsidy effect of $.040***$, significant at the 1% level.

The tax credit effects are calculated in the same way, taking into account the tax credit interaction terms: $lincom (tax_invSE + invsepcc) + tax_invSEincremental*.4166667 +$

$tax_invSEgrowth_related*.3076389 + (tax_invSEhigh_tech + invSEhigh_tech)*.2617934 +$
 $(tax_invSEmanufacturing + invSEmanufacturing)*.4313492 +$
 $(tax_invSEt_start_1996+invSEt_start_1996)*.725 + (tax_invSEpanel+invSEpanel)*.5902778 +$
 $(tax_invSEmicro_smes + invSEmicro_smes)*.0426796 +$
 $(tax_invSEshort_run+invSEshort_run)*.1180556 + (tax_invSEdid + invSEdid)*.1944444 +$
 $(tax_invSEiv + invSEiv)*.2291667 + (tax_invSErdperformersonly +$
 $invSErdperformersonly)*.3666667 + (tax_invSEdeveloping + invSEdeveloping)*.3421053 +$
 $(tax_invSEbinary+invSEbinary)*.6597222 +$
 $(tax_invSEno_control_endogeneity+invSEno_control_endogeneity)*.2666667 +$
 $(tax_invSEtax_domination_1+invSEtax_domination_1)*.45 +$
 $(tax_invSEsub_domination_1+invSEsub_domination_1)*.4210526 + invSEstudy_24*.0833333 +$
 $invSEstudy_26*.0833333 + invSEstudy_28*.0833333 + invSEstudy_29*.0833333 +$
 $invSEstudy_31*.0833333 + invSEstudy_33*.0833333$. Note that the first WLS model is study
 weighted, giving each study equal weight in the regression. Hence, the weights on the tax credit study
 effects – $invSEstudy_24 - invSEstudy_33$ – are the same.

This sums to the average tax credit effect of .067***, significant at the 1% level.

Method of calculating elasticity effects:

With reference to Eq.12:

- For the tax credit literature, the evolution of the authentic effect is derived as $\widehat{\beta}_4 + \widehat{\beta}_5$; while the evolution of publication bias is derived $\widehat{\beta}_6 + \widehat{\beta}_7$. For the subsidy literature, these are directly estimated by, respectively, $\widehat{\beta}_4$ and $\widehat{\beta}_6$.
- The representative elasticity effects for our two literatures are derived as follows. For the subsidy literature, $\widehat{\beta}_0 + \widehat{\beta}_4 * swm_{t_start_1998}^{sub}$. And for the tax credit literature: $(\widehat{\beta}_0 + \widehat{\beta}_3) + (\widehat{\beta}_4 + \widehat{\beta}_5) * swm_{t_start_1998}^{tax}$.
- The publication bias for our two literatures are derived as follows. For the subsidy literature, $\widehat{\beta}_1 + \widehat{\beta}_6 * swm_{yearofpublication_2009}^{sub}$. And for the tax credit literature: $(\widehat{\beta}_1 + \widehat{\beta}_2) + (\widehat{\beta}_6 + \widehat{\beta}_7) * swm_{yearofpublication_2009}^{tax}$.

Online Appendix D. Elasticity regressions

	(1)	(2)	(3)	(4)
VARIABLES	Not Study- Weighted – Default SEs	Study- Weighted – Default SEs	Not Study- Weighted – Default SEs	Study- Weighted – Default SEs
<i>invSE_loglog</i> (inverse SE of the elasticity)	.003** (.001)	.003** (.002)	.007*** (.002)	.007*** (.001)
<i>taxcredit_literature</i> (tax credit dummy)	.38 (.74)	.40 (.83)	1.26 (1.29)	.99 (1.24)
<i>tax_invSE_loglog</i> (inverse SE of the elasticity interacted with tax credit dummy)	.009 (.013)	.004 (.015)	.025 (.017)	.025 (.021)
<i>tax_invSEt_start_1998</i> (start-point of data in 1998 or later interacted with tax credit dummy)			-.048* (.0245)	-.045* (.0256)
<i>invSEt_start_1998</i> (start-point of data in 1998 or later)			.018* (.009)	.015*** (.005)
<i>tax_yearofpublication_2009</i> (year of publication in 2009 or later interacted with tax credit dummy)			.22 (1.50)	.01 (1.42)
<i>yearofpublication_2009</i> (year of publication in 2009 or later)			2.55** (.98)	2.76*** (.62)
<i>_cons</i> (constant)	.73* (.39)	1.12*** (.38)	-1.82* (.95)	-1.56*** (.57)
Number of observations	56	56	56	56
R-squared	.121	.088	.376	.520

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Online Appendix E. The statistical power of the tax credit and subsidy literatures

Ioannidis et al. (2017: 247) note that ‘in many disciplines there has been mounting attention to the issue of statistical power’. Accordingly, we investigate the “median power” for both the tax credit and subsidy literatures as a measure of the typical statistical power of an empirical literature and thus one indicator of research quality.

Ioannidis et al. (2017) draw attention to the use of authentic empirical effects identified by MRA to assess, retrospectively, the statistical power of both individual studies and estimates and, hence, the representative statistical power of empirical literatures. Statistical power is the probability of rejecting a null hypothesis when it is false; hence, the greater the statistical power, the greater the likelihood of minimising “false positives” and detecting ‘a genuine empirical effect’ (Ioannidis et al., 2017: 239). Adequate statistical power is conventionally defined as 80 per cent or more (Cohen, 1965), which means that at least 80 per cent of the distribution of the estimated standardised effect size must lie beyond the critical value for the rejection of the null that the effect size is zero.²⁹

Following Ioannidis et al. (2017), we calculate (i) the “median power” for both the tax credit and subsidy literatures and (ii) the number of estimates in each literature that are adequately powered. For the representative authentic empirical effects (PCCs) estimated by WLS for the tax credit and subsidy literatures respectively (Table 3, Column 1), the levels of median power calculated from the regression sample are 70 per cent and 18 per cent.³⁰ These measures of statistical power are the probabilities that a typical empirical investigation in the tax credit or subsidy literatures can reject a finding when false; conversely, these measures of statistical power may be interpreted as the probabilities of being able to identify tax credit or subsidy effects when they are in the data. Consistent with our median power calculations, we find that the numbers of adequately powered estimates in the tax credit and subsidy regression samples are, respectively, 66 (from six studies) and three (from two studies).

Ioannidis et al. (2017) drew upon 159 MRA databases to discover that the median powers of their respective literatures are between 10% and 18%. From this perspective, while the median statistical power of the tax credit literature approaches the conventional threshold (70% compared to 80%), the subsidy literature has a median statistical power similar to the best performing economic literatures (18%). Yet, such comparisons set the bar low. This MRA suggests that the effect sizes in both literatures are very small. This implies that future studies, if they are to be adequately powered, will require more observations than are typically used in the existing literature.³¹

Small effect sizes combined with the desirability of adequate statistical power bring to the fore trade-offs confronting evaluation research. Administrative datasets typically lack variables required for theory-consistent model specification and so must be linked to other,

²⁹ Tom Stanley, one of the authors of Ioannidis et al. (2017), gives a clear pedagogic explanation of this calculation and of how to supplement MRA by calculating median power:
<https://www.youtube.com/watch?v=9JZ5qqktWYo&feature=youtu.be>

³⁰ Recalculating for the mean of the tax credit and subsidy effects reported in Table 3 (i.e. across the four models), the respective statistical powers are 67 per cent and 17 per cent.

³¹ The (unweighted) median degrees of freedom for the tax credit (subsidy) literature are: 1,367 (677).

less comprehensive datasets, which may reduce an initial sample of millions of firms to one of a few thousand (see, for example, Dechezleprêtre et al., 2016).