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

Chris Dimos  
University of Bath  
School of Management  
c.dimos@bath.ac.uk

Geoff Pugh  
Staffordshire University  
Business School  
g.t.pugh@staffs.ac.uk

Mehtap Hisarciklilar  
Staffordshire University  
Business School  
m.hisarciklilar@staffs.ac.uk

Abstract

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

Name of PhD student: Chris Dimos  
Affiliation: School of Management, University of Bath, Bath, United Kingdom  
Year of PhD enrolment: October 2017  
Expected final date: October 2020  
e-mail address: c.dimos@bath.ac.uk

State-of-the-art

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). The evaluation of the effectiveness of these two instruments is crucial in determining not only whether innovation policy addresses the underlying market failures but also to which extent? if any? additionality in private R&D expenditure is achieved.

Research gap

Although there are large primary literatures that evaluate the impact either of R&D tax credits or of R&D subsidies, none of the studies in either
literature compares the effectiveness of these policy instruments. Accordingly, this study conducts Meta-Regression Analysis (MRA) of the combined literature with the primary purpose of comparing the effectiveness of tax credits and subsidies in promoting private R&D. Beyond the overall effectiveness of the two measures, this study further investigates whether the effectiveness of the two measures varies: a) across different contexts (e.g. firm size and country); and b) across different research practices followed in the literature. Furthermore, the study will investigate whether the effectiveness of each measure increases over time as it is hypothesised by Becker (2015). Finally, the extent to which the primary literatures are susceptible to publication selection bias and whether this bias is time-evolving is examined.

Theoretical arguments

While R&D subsidies target specific projects with high social returns, R&D tax credits are directed to all eligible firms. Due to the mode of implementation of the latter, a prima facie advantage of R&D tax credits is their relative immunity to policy inefficiencies, since they are based upon agents’ optimisation decisions. Conversely, R&D subsidies 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). Accordingly, this and other idiosyncratic characteristics of these two policy instruments suggest the possibility of differential effectiveness.

Method

An integrated database across the two literatures with compatible effect sizes and moderator variables is created. The database includes 540 observations, which are more or less balanced between tax credit and subsidy effects. Because the reported effect sizes are not standard, these will be analysed after transformation into partial correlation coefficients (Doucouliagos and Stanley, 2009; Stanley and Doucouliagos, 2012). The methodology followed is in line with the MRA techniques developed by Stanley and Doucouliagos (2012) which are extended to accommodate the Bayesian Model Averaging technique (Ir?ov? and Havr?nek, 2013). Subsets of both literatures containing small numbers of R&D expenditure effects reported as elasticities are further analysed to provide a quantitative extension of the comparative analysis.

Results

It is found that neither instrument is more effective in promoting private R&D investment than the other. In round terms, an additional $1 of public R&D support of either type induces 7.5 cents of additional private R&D expenditure. However, sources of heterogeneity in the reported effects caution that different contexts may cause differential effectiveness; in particular: subsidies may be the appropriate mode of support for financially constrained firms in late-industrialising or transitional/developing economies; tax credits are less effective for SMEs; and neither tax credits nor subsidies are more or less effective with respect to sectoral and technological heterogeneity. In addition, in relation to tax credits, we find that incremental- and volume-based systems do not yield systematically different effects. Other sources of heterogeneity may help to inform research practices. In particular, approaches to minimising the influence of selection bias reduce reported effect sizes in both literatures: samples of R&D performers only (to minimise differences between treatment and comparison firms); investigating growth effects rather than levels effects; and econometric methods designed to control for both unobservable and observable differences between firms.

In addition, in both literatures, we find that the reported effects are tending to increase over time, as conjectured by Becker (2015). Given that we control for the influence of both publication selection bias and sources of heterogeneity, this evidence identifies authentic changes in policy effects, which may reflect learning by public authorities and/or firms. Finally, both literatures suffer from positive publication selection bias which is increasing over time but this bias is more pronounced in the tax credit literature.

References


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

Chris Dimos\textsuperscript{a,c}, Geoff Pugh\textsuperscript{b,d}, Mehtap Hisarciklilar\textsuperscript{b,e}

\textsuperscript{a} School of Management, University of Bath, Claverton Down, Bath BA2 7AY, UK
\textsuperscript{b} Centre for Applied Business Research (CABR), Staffordshire University, Leek Road, Brindley Building, Stoke-on-Trent ST4 2DF, UK
\textsuperscript{c} c.dimos@bath.ac.uk
\textsuperscript{d} Corresponding author. Tel.: +44 0 1782 294092. g.t.pugh@staffs.ac.uk, geoff@plainfigures.com
\textsuperscript{e} m.hisarciklilar@staffs.ac.uk

Abstract

Although there are large primary literatures that evaluate the effectiveness of either R&D tax credits or R&D subsidies in promoting private R&D, no previous study systematically measures and compares the effectiveness of these two policy instruments. Each of these literatures has been the subject of recent Meta-Regression Analysis (MRA) – Castellacci and Lie, 2015 (tax credits) and Dimos and Pugh, 2016 (subsidies) – but the attention of both studies is confined to their respective primary literatures. This MRA investigates the combined literature to measure and compare the effectiveness of these two instruments. After controlling for publication selection and sources of heterogeneity, we find that both tax credits and subsidies induce small additional private R&D investment and that neither instrument systematically outperforms the other. In both literatures, we find that the authentic effects – i.e. beyond publication bias and sources of heterogeneity – are tending to increase over time. Moreover, heterogeneity in the reported effects arise from both different contexts (e.g. firm size and country) and different research practices.

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

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. However, 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,\(^1\) while Becker (2015) has provided a narrative review of the empirical evidence from both literatures. Both MRA studies find that, in their respective literatures, the R&D support effects are exaggerated by publication selection bias. In addition, Castellacci and Lie (2015) highlight that the tax credit effects are sectorally differentiated, whereas Dimos and Pugh (2016) focus on deriving a representative effect from the subsidy literature and find that (i) R&D subsidies do not crowd out private R&D investment but achieve only limited additionality and (ii) both sample and methodology heterogeneities can explain variation in reported subsidy effects. Becker (2015) contributes, in particular, the conjecture that, over time, both tax credits and subsidies are becoming more effective in promoting private R&D investment.\(^2\) This study applies MRA to both literatures jointly to determine the differential 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 in each literature.

We find that both tax credits and subsidies give rise to additionality of a similarly small size. Yet, 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 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.

2. Context and theoretical background

R&D tax credits and R&D subsidies are not perfect substitutes, although both are expected to lead to an increase in private R&D investment (David et al., 2000). 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 directed to all eligible firms by taking into account either their overall R&D spending (volume-based system) or their excess R&D

\(^1\) Unfortunately, methodological differences between these two MRAs preclude using them to compare the effectiveness of the two instruments.

\(^2\) 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’.
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. This renders tax credits 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 on R&D tax incentives (mainly tax credits) and direct R&D funding (mainly subsidies) expressed as a percentage of GDP. While both overall public R&D support and the proportion between the two measures vary across countries, in most cases governments use both instruments.3

Figure 1. Direct government funding of business R&D and tax incentives for R&D, 2013 (as a percentage of GDP)

*BERD (Business Expenditure on R&D)

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

---

3 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).
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 some guidance on the relative effectiveness of tax credits and
subsidies in particular 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 arms-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, according to
Mohnen n.d., 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); 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). An implication for the analysis to follow is that while subsidy effects will be tested by
conventional two-tail tests (reflecting a range from negative to positive) tax-credit effects can
be evaluated by one-tail tests (reflecting a truncated range of zero and positive). However,
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

To investigate the relative effectiveness of R&D tax credits and R&D subsidies, we require comparable effect sizes across the two literatures. Only one segment from each literature reports comparable effect sizes: i.e. estimates of the rate at which private R&D expenditure increases due to the introduction of, respectively, fiscal incentives or subsidies.\(^4\) Figure 2 shows the coverage of our MRA database as the union of these two segments: the “Additionality Ratio” segment of the tax credit literature; and the “Private R&D” segment of the subsidy literature. Castellacci and Lie (2015:819) and Dimos and Pugh (2016:800) explain both these and the other segments of the respective literatures, which are all depicted in Figure 2.

Figure 2. The MRA database in relation to the segments of the tax credit and subsidy literatures

The resulting MRA database comprises 251 tax credit effects from 12 studies and 289 subsidy effects from 23 studies.\(^5\) Our database thus includes nearly half of the estimated effects reported in the two literatures in or after 2000. In our database, the typical tax credit

\(^4\) The effects of both types of public support may be estimated within either a matching or a regression framework. Both yield three possible effects: statistically significant and positive – 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.

\(^5\) The list of coded studies is available at: www.researchdata.bath.ac.uk/
(subsidy) study reports 21 (13) effect sizes with the median being 15.5 (12); across these studies, the minimum number is 1 (1) and the maximum 72 (34).

3.2 Effect sizes and public support outcomes

Figure 3 depicts the reported effects from both literatures. In accord with theory, the subsidy literature reveals all three possible outcomes: 6.6% (19 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. Crowding out is precluded by theory, although 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). The absence of more perverse results is possibly an indicator of publication selection in this literature, a theme to be investigated below. Overall, although additionality is the dominant finding, a substantial minority of estimates suggest no effect from public support of private R&D.

![Figure 3. Reported effects from public support of private R&D](image)

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:

$$ PCC = \frac{t}{\sqrt{t^2 + df}} $$  \tag{1} $$

$$ SE_{PCC} = \sqrt{\frac{1 - PCC^2}{df}} $$  \tag{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.

### 3.4 Descriptive statistics

Funnel plots of the tax credit and subsidy effect sizes are presented in Figure 4. By plotting the transformed effect size ($PCC$) against its precision (the inverse of the $SE_{PCC}$) (Stanley and Doucouliagos, 2012:53-60) we gain initial insights into the literature, in particular with respect to potential heterogeneities and publication selection bias. The funnel plots for the combined sample of R&D “support” effects (omitted for reasons of space) as well as the separate plots reported in Figure 4 suggest that we are dealing with a coherent literature. In both the tax credit and the subsidy literatures the “inverted funnel” indicates roughly similar uni-modal distributions. This suggests a reasonably homogeneous literature, whereas a multi-modal distribution would question the decision to combine the literature as a coherent object of analysis.

The plots are also informative about publication selection in the two literatures. In accord with sampling theory, as precision levels rise both plots show funnel “stems” rising towards similar small, positive PCCs with lower dispersion. In both literatures, this suggests a small average additionality effect. However, as we move from higher levels of precision towards lower levels of precision the dispersion widens, but it does not widen symmetrically. In the meta-analysis and MRA literatures, skewed distribution at lower levels of precision is interpreted as the trace of publication selection in the literature. This arises from individual decisions to report estimates only of a certain sign and of a sufficient size to compensate for low precision, thereby yielding theory-consistent “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 – i.e. “beyond” – publication bias (Stanley, 2005). In the literature under consideration, the funnel plot for the tax credit sample (Panel a) is somewhat skewed to the right, indicative of positive publication bias. In contrast, the funnel plot for the subsidy sample (Panel b) seems to be more or less symmetric around a small positive PCC, indicative of a small average additionality effect not apparently influenced by publication bias. These preliminary observations on publication bias are confirmed by the MRA to follow.

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 weight each estimated effect by the inverse number of estimates reported in the source study. In the study-weighted funnel plots (panels c and d) the size of the dots represent the weight assigned to the corresponding estimate (the larger the weight the larger the dot). Our impressions from the study-weighted funnel plots confirm those from the study-unweighted plots; namely, small positive effects (additionality) in both literatures with an indication of positive publication bias in the tax credit literature.

Figure 4. Funnel plot (study-unweighted and study-weighted effect sizes)

![Funnel plot](image)

Source: authors.

Overall, the evidence from this preliminary investigation suggests that (i) both tax credit and subsidy effects are positive although (ii) the respective literatures may be differentially influenced by publication selection, with strong indications of positive publication bias in the tax credit literature but, at most, weak indications in the subsidy literature.
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 bias and sources of heterogeneity, where these may be sample characteristics – e.g. firm, industry and/or country – or different research practices.

4.1 Model specification and estimation strategy

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):

\[ PCC_i = \beta_0 + \beta_1 SE_i^{PCC} + \epsilon_i \]  

(3)

where \( i = 1, \ldots, n \) indexes the \( n \) individual estimates reported in the primary literature, \( SE_i^{PCC} \) denotes the standard error of the \( i^{th} PCC \) and \( \epsilon_i \) is the usual regression error. The statistical significance of \( \beta_1 \) 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).

To correct for heteroskedasticity, which is a characteristic of Eq. (3) because the variance of the \( PCC_i \) (and, thus, the variance of \( \epsilon_i \)) is not constant, Weighted Least Squares (WLS) estimation is employed. WLS estimation is implemented by dividing Eq. (3) by the standard error of \( PCC_i \left( SE_i^{PCC} \right) \) (Stanley and Doucouliagos, 2012:61), which not only addresses heteroskedasticity but also implements precision weighting (i.e. more precise estimates are given greater weight):

\[ \frac{PCC_i}{SE_i^{PCC}} = t_i = \beta_1 + \beta_0 \left( \frac{1}{SE_i^{PCC}} \right) + \nu_i \]  

(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), \( \nu_i = \epsilon_i / SE_i^{PCC} \) is the error term corrected for heteroskedasticity, and \( 1/SE_i^{PCC}, \) the inverse of the \( SE \) on the \( PCC, \) is the precision term. In the transformation of Eq. (3) into Eq. (4) the coefficients \( \beta_0 \) and \( \beta_1 \) change place but their interpretation is unchanged. To avoid confusion, we recall from Section 3 that we also weight our regressions by the inverse number of estimates reported in each study, so that results are not unduly influenced by studies reporting large numbers of estimates. Accordingly, our estimates reflect two types of weighting: precision weighting; and study weighting.

Our estimation strategy is not to estimate tax credit and subsidy models separately, by dividing our MRA database into its component literatures, but instead to pool them and use the full resources of our data to exploit efficiency gains (Greene, 1993:236). 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 \((D_{i}^{\text{tax}})\) – 1 (0) denotes estimates from the tax credit (subsidy) literature – and an interaction term between \(D_{i}^{\text{tax}}\) and \(1/SE_{i}^{\text{PCC}}\). In this procedure, subsidy effects are estimated directly but tax credit effects are derived post estimation as linear combinations of estimated coefficients. This yields the same point estimates as separate sub-sample regression but with slightly smaller standard errors:

\[
t_i = \beta_1 + \beta_0\left(\frac{1}{SE_{i}^{\text{PCC}}}\right) + \beta_2 D_{i}^{\text{tax}} + \beta_3 \left(\frac{1}{SE_{i}^{\text{PCC}}} D_{i}^{\text{tax}}\right) + \nu_i
\]

(5)

where \(\beta_1\) (the intercept) controls for and measures publication bias in the subsidy literature; \(\beta_2\) measures the effect of \(D_{i}^{\text{tax}}\), 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, \(\beta_0\) measures the authentic effect for the subsidy literature, \(\beta_3\) 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” \(Z_i\), i.e. dummy variables that capture study-related characteristics that condition the effect sizes reported in the primary literature(s). Eq. (6) augments Eq. (5) with two sets of variables:

(a) moderator variables interacted with the precision effect \(Z_{mi}/SE_{i}^{\text{PCC}}\); and
(b) moderator variables interacted with the precision effect and the tax credit dummy \((Z_{mi}/SE_{i}^{\text{PCC}})D_{i}^{\text{tax}}\), to allow for differential moderator effects across the two literatures.

\[
t_i = \beta_1 + \beta_2 D_{i}^{\text{tax}} + \beta_0 \left(\frac{1}{SE_{i}^{\text{PCC}}}\right) + \beta_3 \left(\frac{1}{SE_{i}^{\text{PCC}}} D_{i}^{\text{tax}}\right) + \sum_m \beta_{m1} \left(\frac{Z_{mi}}{SE_{i}^{\text{PCC}}}\right) + \sum_m \beta_{m2} \left(\frac{Z_{mi}}{SE_{i}^{\text{PCC}}} D_{i}^{\text{tax}}\right) + \nu_i
\]

(6)

where \(m\) indexes moderator variables. The estimates of \(\beta_{m1}\) measure the effect of each of the \(m\) moderator variables conditioning estimated effects in the subsidy literature; the estimates of \(\beta_{m2}\) 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 \(\beta_0\), the estimated coefficient on the precision term (1/\(SE_{i}^{\text{PCC}}\), and the \(\beta_{m1}\), the estimated coefficients on each of the interacted moderator variables \((Z_{mi}/SE_{i}^{\text{PCC}})\) 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 downweighted 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). 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: \( \beta_0 \) and \( \beta_3 \), the estimated coefficient on the differential precision term for the tax credit literature \( (1/SE_{i PCC}^t)D_{i tax}^t \); and the sum of each of the \( \beta_{m1} \) and each of the corresponding \( \beta_{m2} \), the estimated coefficients on the differential tax credit moderator variables \( (Z_{mi}/SE_{i PCC}^t)D_{i tax}^t \), weighted by the respective study-weighted means.

Our baseline WLS model is the standard but not the only approach 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). First, we estimate a variant on fixed-effects estimation. This takes into account the 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 two ways: by reporting weighted estimates giving each study equal influence; and by reporting cluster-robust SEs, which are robust to arbitrary patterns of dependence among the residuals from estimates from the same study. Here, in addition, we augment Eq. (6) to control for the moderating effects of each specific study on the estimated authentic effects of R&D support:

\[
t_{is} = \beta_0 + D_{i tax}^t + \beta_1 \left( \frac{1}{SE_{is PCC}^t} \right) + \beta_2 \left( \frac{1}{SE_{PCC}^t} \right) + \sum_m \beta_{m1} \left( \frac{Z_{mis}}{SE_{is PCC}^t} \right)
+ \sum_m \beta_{m2} \left( \frac{D_{i tax}^t Z_{mis}}{SE_{is PCC}^t} \right) + \sum_s \left( \frac{\mu_s}{SE_{is PCC}^t} \right) + \nu_{is}
\]

where \( s \) indexes the studies in the primary literature. The difference between this fixed-effects model and the WLS model in Eq. (6) is the study-specific effects \( \mu_s \) (a fixed effect – 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. \( \mu_s \left(1/SE_{is PCC}^t\right) \).

In addition, we implement two more approaches to the estimation of our baseline model. First, 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). Second, 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). Robust regression weighting precludes the additional use of study weighting. Hence, we estimate three study-weighted models – WLS, FE and BMA – and one study-unweighted model.

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 apply all four of our estimation approaches to both (i) a model specified only with the Z

---

6 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.
moderators and (ii) an alternative specification in which the $Z$ moderators are augmented by a single $K$ moderator – “Year of publication” – which allows publication bias to evolve over time. Among our $Z$ moderator variables is the “Starting point of data”, which captures potential time variation in the authentic effects. In the augmented specification, 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. Apart from this single $K$ moderator in the augmented specification, we focus on $Z$ moderators, because our primary purpose is to estimate and compare the authentic effects – if any – from the tax credit and subsidy literatures. Accordingly, our objective is to measure and control for rather than to analyse publication bias.

4.2 Results and discussion

We do not report our estimates from both models across all four approaches to estimation. Instead, we report the main findings of our study in Table 1. 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 eight models estimated.

The regression coefficients in each multiple MRA regression are estimated net of publication bias because they are estimated conditional on the regression constant and the year of publication (the associated $K$ moderator variable), which captures publication bias. In turn, the authentic effect size revealed by the literature is calculated from these estimated coefficients to not only filter out publication bias but also to reflect the heterogeneities in the literature. The authentic effect size is obtained from each regression by summing the estimated coefficient on the precision term with the coefficients on the $Z$ moderator variables, each of which is weighted by the frequency with which it appears in the literature. (The weighting ensures that a $Z$ moderator reported in few regressions and/or in studies reporting many estimates will be down-weighted compared to a $Z$ moderator reported in many regressions and/or in studies reporting few estimates). (The method of derivation is explained in detail following Eq.6 above; see online Appendix A for indicative examples of the syntax used to calculate the authentic effects and publication bias.8).

To aid interpretation of the estimates reported in Table 1, 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 ≤ PCC ≤ 0.33); or “large” (PCC > 0.33).

---

7 These results are available online at: www.researchdata.bath.ac.uk/
8 Online appendices are available at: www.researchdata.bath.ac.uk/
We provide a qualitative overview of our findings in Table 2, which reports findings simply as positive and statistically significant (+), negative and significant (-), or as statistically insignificant (0).\footnote{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.}
<table>
<thead>
<tr>
<th></th>
<th>WLS (weighted – cluster robust)</th>
<th>ROBUST REGRESSION</th>
<th>FE (weighted – cluster robust)</th>
<th>BMA (weighted)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline model</td>
<td>Augmented model</td>
<td>Baseline model</td>
<td>Augmented model</td>
</tr>
<tr>
<td>Average tax credit effect</td>
<td>.017* (.012)</td>
<td>.013 (.014)</td>
<td>.064*** (.10)</td>
<td>.044*** (.10)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.057*** (.10)</td>
<td>.074*** (.08)</td>
</tr>
<tr>
<td>Average subsidy effect</td>
<td>.051* (.026)</td>
<td>.059*** (.20)</td>
<td>.026*** (.10)</td>
<td>.046*** (.08)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.057*** (.16)</td>
<td>.041*** (.09)</td>
</tr>
<tr>
<td>(Tax credit – subsidy) effect</td>
<td>-.034 (.029)</td>
<td>-.046* (.25)</td>
<td>.038*** (.14)</td>
<td>-.001 (.13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.0002 (.192)</td>
<td>.033*** (.12)</td>
</tr>
<tr>
<td>(Tax credit – subsidy) PB</td>
<td>1.77*** (.50)</td>
<td>1.92*** (.55)</td>
<td>.05 (.32)</td>
<td>.74** (.34)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.40 (.36)</td>
<td>-.23 (.30)</td>
</tr>
<tr>
<td>Tax credit PB</td>
<td>.34 (.86)</td>
<td>.21 (.64)</td>
<td>.94*** (.26)</td>
<td>.35 (.22)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.06 (.42)</td>
<td>.46** (.22)</td>
</tr>
<tr>
<td>Subsidy PB</td>
<td>1.43 (.99)</td>
<td>1.71** (.84)</td>
<td>-.89** (.42)</td>
<td>.38 (.40)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.34 (.55)</td>
<td>-.69* (.37)</td>
</tr>
<tr>
<td>(Tax credit – subsidy) PB</td>
<td>1.65† (.45)</td>
<td>2.02† (.47)</td>
<td>1.43 (.99)</td>
<td>1.71** (.84)</td>
</tr>
<tr>
<td>Average subsidy effect</td>
<td>.051* (.026)</td>
<td>.059*** (.20)</td>
<td>.026*** (.10)</td>
<td>.046*** (.08)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.057*** (.16)</td>
<td>.041*** (.09)</td>
</tr>
<tr>
<td>(Tax credit – subsidy) effect</td>
<td>-.034 (.029)</td>
<td>-.046* (.25)</td>
<td>.038*** (.14)</td>
<td>-.001 (.13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.0002 (.192)</td>
<td>.033*** (.12)</td>
</tr>
<tr>
<td>(Tax credit – subsidy) PB</td>
<td>1.77*** (.50)</td>
<td>1.92*** (.55)</td>
<td>.05 (.32)</td>
<td>.74** (.34)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.40 (.36)</td>
<td>-.23 (.30)</td>
</tr>
<tr>
<td>Tax credit PB</td>
<td>.34 (.86)</td>
<td>.21 (.64)</td>
<td>.94*** (.26)</td>
<td>.35 (.22)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.06 (.42)</td>
<td>.46** (.22)</td>
</tr>
<tr>
<td>Subsidy PB</td>
<td>1.43 (.99)</td>
<td>1.71** (.84)</td>
<td>-.89** (.42)</td>
<td>.38 (.40)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.34 (.55)</td>
<td>-.69* (.37)</td>
</tr>
<tr>
<td>(Tax credit – subsidy) PB</td>
<td>1.65† (.45)</td>
<td>2.02† (.47)</td>
<td>1.43 (.99)</td>
<td>1.71** (.84)</td>
</tr>
</tbody>
</table>

* Statistically significant at the 10% level in a one-tail test.
† “Significant” in the BMA sense of ‘robustly correlated with the dependent variable’ (De Luca and Magnus, 2011:533)
Cluster-robust standard errors in parentheses: *** p<0.01; ** p<0.05; * p<0.1
Table 2. Summary of results: sources of heterogeneity; publication bias; and authentic effects

<table>
<thead>
<tr>
<th>Research Practices:</th>
<th>Tax Credit (TC)</th>
<th>Subsidy</th>
<th>Comparison (TC – Subsidy)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WLS</td>
<td>RR</td>
<td>FE</td>
</tr>
<tr>
<td>A. Moderator variables (MVs)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contextual:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incremental-based system</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>R&amp;D performers only</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Micro and SMEs only</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Hi-tech only</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Manufacturing sector only</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Developing economies</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B. Publication Bias and Authentic R&amp;D Support Effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average publication bias</td>
<td>+</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Robustness check: separate-sample estimation</td>
<td>+</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Publication bias evolution: Year of publication (MV)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Authentic effects</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Robustness check: separate-sample estimation</td>
<td>+</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Authentic effect evolution: Start-point of data (MV)</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

*One-tail test. Key: +’ve / -’ve indicate positive/negative and statistically significant at the 10% level (WLS, RR and FE); t-statistic > 1 in absolute value (BMA). 0 indicates not statistically different from zero (WLS, RR and FE); t-statistic < 1 in absolute value (BMA).
The qualitative overview of our estimates presented in Table 2 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. 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.

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. (2012:24): ‘There is hardly any pattern observable in the results of the evaluations of incremental or volume-based tax incentive schemes.’ In common with Castellacci and Lie (2015) we reach a similar conclusion: two from eight estimates suggest that incremental schemes yield smaller effects on private R&D than do volume or mixed incremental and volume schemes; two suggest larger effects; and four estimates are not statistically significant.

- Studies evaluating R&D support effects on homogeneous samples of “R&D performers only” report smaller effects: whenever the estimate is statistically significant, it is negative (five from eight in the tax credit literature – tax_invSErdperformersonly+ invSErdperformersonly; four from eight in the subsidy literature – invSErdperformersonly). We illustrate the quantitative effect of sample homogeneity with the robust regression estimates from the baseline model: other factors held constant, the tax credit PCC is reduced by .040; and the subsidy PCC by .029. These findings are consistent with Dimos and Pugh (2016) for the subsidy literature, although this effect was not investigated by Castellacci and Lie (2015) for the tax credit literature. A more homogeneous sample reduces differences between treated and untreated firms which, in turn, may reduce selection bias. 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).

- Contrary to Castellacci and Lie (2015) who find that tax credits are more effective in promoting R&D investment by SMEs than by larger firms, we report four from eight estimates suggesting that tax credits are less effective for micro and SMEs than for large firms (the four non-significant estimates are likewise all negative) (tax_invSEmicro_smes + invSEmicro_smes), although these effects are all “small”. This evidence is complementary to Busom et al. (2014:571) who hypothesise that ‘the exclusive use of tax credits will be negatively correlated with high financing constraints’. This applies with particular force to SMEs, because these firms ‘are more likely to face’ financing constraints (reflecting their ‘lack of collateral’, ‘lack of reputation’ – especially if young – and ‘sunk costs’) (Busom et al., 2014:579 and 581). Although the effects estimated by Busom et al. (2014:585) are not supportive of their hypothesised differential propensity of SMEs and large firms to use tax credits, the reduced effectiveness of tax credits for SMEs revealed by the present study is in line with their hypothesised effect. In comparison with
the single significantly negative SME effect reported by Dimos and Pugh (2016), we find no evidence to distinguish subsidy effectiveness by firm size (invSEmicro_smes) (all eight estimates suggest no significant difference). This is consistent with Busom et al. (2014) who also find no evidence that subsidies are more effective for SMEs than for large firms. Together, our evidence from both literatures is consistent with the conclusion of Busom et al. (2014:587) that tax credits and subsidies are ‘not substitutes’ for addressing the financial constraints and other obstacles to R&D faced by SMEs.

- Whereas Castellacci and Lie (2015) find that tax credits are less effective in promoting R&D investment by hi-tech firms relative to firms in lower technology categories, we find no evidence to suggest this (tax_invSEhigh_tech + invSEhigh_tech) (all eight estimates are statistically insignificant). Conversely, Dimos and Pugh (2016) report lower effectiveness of subsidies for hi-tech firms, with which the evidence from the present study is weakly consistent (one positive estimate, two negative and the rest not significant).

- Similar to Castellacci and Lie (2015), we find little evidence of differential effectiveness of tax credits by broad sector (manufacturing with respect either to services or to manufacturing and services jointly) (seven statistically insignificant effects from eight; tax_invSEmanufacturing + invSEmanufacturing). And, similar to Dimos and Pugh (2016), the evidence for subsidies (invSEmanufacturing) is too mixed to draw any conclusion.

- Tax credits are less effective in developing economies than in developed economies (four from eight estimates of the effect of tax_invSEdeveloping + invSEdeveloping are negative and statistically significant; the other four are negative but insignificant). Conversely, consistent with Dimos and Pugh (2016), subsidies appear to be more effective in developing than in developed economies (invSEdeveloping; six positive effects, two negative). In addition, seven from eight statistically significant comparisons suggest that in developing economies subsidy effects are greater than tax credit effects (tax_invSEdeveloping). Of the three studies (33 observations) on subsidies from developing economies one is from Turkey (Ozcelik and Taymaz, 2008; eight observations) and two from eastern Germany (Alecke et al., 2012 and Almus and Czarnitzki, 2003; 25 observations), which both studies classify as developing. Ozcelik and Taymaz (2008) hypothesise that R&D subsidies could be more effective than R&D tax incentives in developing/late-industrialising economies. For eastern Germany, both studies draw on data close to the onset and mid-point (1995-97 and 2003, respectively) of eastern Germany’s industrial restructuring during which tax credits and investment subsidies of all types were strategic elements (Lange and Pugh, 1998: 68-69 and 77). Given that subsidies may be more effective in promoting R&D investment by financially constrained firms (Busom et al., 2014; and, in particular, for transition economies, Becker, 2013:26), the prevailing ‘lack of profitability’ of the corporate sector in eastern Germany (Lange and Pugh, 1998:132 and 148-49) could explain why subsidies proved more effective than tax credits. In the absence of more comprehensive evidence from developing economies, we cannot generalise these findings.

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

- Whereas Marino et al. (2016:14) find a strong additionality 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 uniformly endorse this conclusion: the eight significantly
negative estimates confirm that R&D expenditure growth effects tend to be smaller than R&D expenditure (levels) effects.\textsuperscript{10} 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: two from eight tax credit effects are significantly positive, with the rest not significant ($\text{tax\_invSE\_no\_control\_endogeneity} + \text{invSE\_no\_control\_endogeneity}$); while this positive effect is more pronounced in the subsidy literature (seven from eight estimates; $\text{invSE\_no\_control\_endogeneity}$). The apparently weaker results for tax credits may be attributed to delivery at “arm’s length”, therefore not being susceptible to selection by programme managers. We conjecture that while subsidies are subject to both selection by programme managers and self-selection, for tax credits only firms’ self-selection is at work.

- We also find that the effects of estimation approaches known to be effective in controlling for both observable and unobservable sources of endogeneity – difference-in-differences (DiD) and instrumental variable estimators (IV) – tend either to reduce the effects reported in both primary literatures or to have no discernible effect compared to other estimation approaches. 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: homogeneous samples of R&D performers; specification of R&D expenditure in growth rather than in levels; estimation approaches that do not control for endogeneity (in comparison with all approaches that do); and 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; Czarnitski et al., 2011; Yang, et al., 2012).

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

- There is weak evidence that the use of panel data has a positive effect in both literatures (in both cases, two from eight estimates), 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’.

\textsuperscript{10} All 68 estimates of growth effects – from 6 studies – are from the tax credit literature.
• The moderator “Short run” captures the effect of reporting estimated coefficients from dynamic panel models which, because these are short-run or impact effects, should be lower than the effects reported from static models. Although only four estimates are statistically significant, these are all negative and thus consistent with this hypothesised effect.

• There is some discussion of the effect of using binary indicators for R&D support rather than continuous data on the value of support. The evidence from our MRA does not give reliable support to arguments favouring one rather than the other in deriving the overall tax credit or subsidy effects (across our 16 estimates, half are statistically insignificant, four are negative and four positive). While Castellacci and Lie (2015) do not investigate this effect, Dimos and Pugh (2016) report no systematic effect for that part of the subsidy literature investigated in the present study.11

In Table 2, Panel B, we report evidence that is consistent with previous indications from funnel plots. Positive publication bias (PB) is found in five of the estimates from the tax credit literature and three from the subsidy literature, with the rest being insignificant. Moreover, four from five significant estimates indicate either “substantial” (1 ≤ |PB| ≤ 2) or “severe” (|PB|>2) publication bias in the tax credit literature, while all three significant estimates indicate “little to modest” publication bias (|PB|<1) in the subsidy literature (Doucouliagos and Stanley, 2013). We also find weak evidence that publication bias has evolved in the tax credit literature (two significantly positive estimates and one significantly negative; tax_yearofpublication_2008+yearofpublication_2008) but uniform evidence of increasing publication bias over time in the subsidy literature (yearofpublication_2008). We offer a tentative explanation of the differential evolution of publication bias in the two literatures. In the tax credit literature, the preclusion of negative – crowding out – effects by theory (see Section 2 above) may have been an overriding and thus permanent source of publication selection from the outset. In contrast, in the subsidy literature, theory is consistent with both negative and positive effects and so was no obstacle to the well-known “decline effect”, whereby many ‘scientifically discovered effects published in the literature seem to diminish over time’ (Schooler, 2011:437). 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 the tax credit literature, the authentic effects are significantly positive in five cases; and in the subsidy literature in all cases. Although Table 2 presents strong evidence that both tax credit and subsidy effects typically increase R&D, these effects are small. Measured as PCCs, the significant overall authentic tax credit effects vary between .017 (in a one-tail test) and .074, while the subsidy effects vary between .026 and .064. There are two striking features of these results: first, we discern no systematic

---

11 There are advantages in using a continuous over a binary measurement of public support receipt. Continuous data enable identification of not only non-linear effects in both literatures but also different degrees of crowding out for the subsidy literature (Dimos and Pugh, 2016).
difference between tax credit and subsidy effects (three of the comparisons are not statistically significant, while two are positive and three negative);\textsuperscript{12} second, consistent with similar strength of association between both types of support and firms’ private R&D, these effects – with one borderline exception – are “small”, according to the guidelines introduced above.

Finally, Becker’s (2015) conjecture that reported R&D support effects are tending to grow over time is supported by the significantly positive coefficients on our “time effect” moderator, which divides studies in the primary literature according to the starting year of their datasets: by four from eight estimates for the tax credit literature (tax\_invSEt\_start\_1996 + invSEt\_start\_1996); and by three from eight for the subsidy literature (invSEt\_start\_1996). Conversely, no contradictory evidence is found (i.e. no significantly negative estimates). Given that we take into account changes in publication practices over time, this evidence identifies authentic changes in policy effects. Hence, it becomes plausible to support Becker’s conjecture by attributing increasing R&D policy effects to learning by public authorities and/or firms and, in particular, to “intra-organisational learning” – essentially learning-by-doing – in the context of innovation policy (Nauwelaers and Wintjes, 2008).

5. Conclusion

This study uses Meta-Regression Analysis (MRA) to compare the effectiveness of R&D tax credits and R&D subsidies in promoting private R&D investment. We find that both R&D tax credits and R&D subsidies typically give rise to additionality. Yet, the additional private R&D investment induced by either tax credits or subsidies is “small”. MRA yields statistically significant estimates of the authentic PCC between public R&D support and private R&D expenditure varying between .017 and .074 for the tax credit literature and between .026 and .064 for the subsidy literature. Our findings uniformly suggest that both measures are similarly effective and yield input additionality. In addition, in both literatures, we find that the reported effects are tending to increase over time, as conjectured by Becker (2015). Given that we control for the influence of both publication selection bias and sources of heterogeneity, this evidence identifies authentic changes in policy effects, which may reflect learning by public authorities and/or firms.

In all of our estimates, we control for publication bias. In estimating authentic effects for the two literatures, we find more and stronger evidence for publication bias in the tax credit literature. This may reflect the theoretical presumption that negative tax credit effects are precluded, which inclines researchers not to report such effects. A particular contribution of this study is to partition the tendency for effect sizes to increase over time into genuine increases in authentic effects and increases in publication bias. PCC effects uniformly reveal positive evolution of publication bias in the subsidy literature. The declining incidence of crowding out and correspondingly increasing incidence of additionality effects in the subsidy literature is consistent with the widely observed “decline effect” in scientific literature. In the tax credit literature, the evidence for a positive evolution is weaker, perhaps reflecting the

\textsuperscript{12} The difference between the authentic effects is the difference between the linear combinations from which each effect is derived.
stabilising effect of the theoretical preclusion of crowding out. We conclude that, after taking account of increasing preference for positive findings in the research community, the increase over time of reported effects is genuine.

Although we find that both tax credits and subsidies are effective in promoting private R&D investment, and that neither systematically outperforms the other, these two types of support are not always substitutes. Sources of heterogeneity in the reported effects caution that different contexts may cause differential effectiveness; in particular: subsidies may be the appropriate mode of support for financially constrained firms in late-industrialising or transitional/developing economies; tax credits are less effective for SMEs; and neither tax credits nor subsidies are more or less effective with respect to sectoral and technological heterogeneity. In addition, in relation to tax credits, we find that incremental- and volume-based systems do not yield systematically different effects. Other sources of heterogeneity may help to inform research practices. In particular, approaches to minimising the influence of selection bias reduce reported effect sizes in both literatures: samples of R&D performers only (to minimise differences between treatment and comparison firms); investigating growth effects rather than levels effects; and econometric methods designed to control for both unobservable and observable differences between firms.

References


13 This is consistent with Busom et al. (2014: 590): ‘… R&D tax credits and R&D subsidies do not appear to be equivalent tools for SMEs.’


