**The effectiveness of R&D subsidies: a meta-regression analysis of the evaluation literature**

**Christos Dimos a**

a Centre for Applied Business Research (CABR), Faculty of Business, Education and Law, StaffordshireUniversity, Leek Road, Brindley Building, Stoke-on-Trent, ST4 2DF, UK. e-mail: [christos.dimos@staffs.ac.uk](mailto:christos.dimos@staffs.ac.uk) , [chrisdimos@hotmail.com](mailto:chrisdimos@hotmail.com)

**Geoff Pugh b, \***

b Centre for Applied Business Research (CABR), Faculty of Business, Education and Law, StaffordshireUniversity, Leek Road, Brindley Building, Stoke-on-Trent, ST4 2DF, UK. e-mail: [g.t.pugh@staffs.ac.uk](file:///E:\2_M%20R%20A%20%20%20P%20A%20P%20E%20R\g.t.pugh@staffs.ac.uk)

**ABSTRACT**

Widespread and increasing public subsidy for research and development (R&D) has given rise to a large and growing number of evaluation studies. While economic theory identifies market failures that justify public support, theory also suggests reasons why returns might be disappointing. Similarly, the empirical literature investigated – 52 micro-level studies published since 2000 on either input or output R&D – reports a wide range of findings. The lack of conclusiveness both of theory and of the evaluation literature motivate this Meta-Regression Analysis (MRA). This study contributes to policy debate by identifying a representative subsidy effect: after controlling for publication selection bias and for a wide range of sample and study heterogeneities, MRA findings reject crowding out of private investment by public subsidy but reveal no evidence of substantial additionality. In addition, among the research practices explaining the heterogeneous effects reported in this literature, those related to the treatment of unobservable firm heterogeneity are particularly important.

**Keywords**: R&D subsidies; evaluation; meta-regression analysis; publication bias; additionality; crowding out

\* CorrespondingAuthor: Professor Geoff Pugh, Director of the Centre for Applied Business Research (CABR), Faculty of Business, Education and Law, StaffordshireUniversity, Leek Road, Brindley Building, Stoke-on-Trent, ST4 2DF, UK. +44 (0)1782 294092 e-mail: [g.t.pugh@staffs.ac.uk](file:///E:\2_M%20R%20A%20%20%20P%20A%20P%20E%20R\g.t.pugh@staffs.ac.uk)

# Introduction

In the Schumpeterian tradition, innovation determines growth through “creative destruction” (Schumpeter, 1942). Similarly, neoclassical growth theory identified technical progress as the only source of sustainable per capita income growth (Solow, 1956). Yet, according to other developments in neoclassical theory in the 1950s and 1960s (Nelson, 1959; Arrow, 1962; Usher, 1964), firms tend to underinvest in innovation due to market failures. Influenced by these theories, policy makers embraced innovation as the means to achieve higher and sustainable growth rates and, correspondingly, governments introduced a variety of policy instruments to promote private innovation, including public subsidies – grants and/or (low-interest) loans – to incentivise private firms to perform research and development (R&D). This study is a contribution to assessing the effectiveness of R&D subsidies**.**[[1]](#footnote-1)

Public subsidies constitute a direct support measure in contrast to indirect fiscal support for R&D (e.g. tax credits) and are widely implemented by public authorities. Figure 1 below indicates the share of product and/or process innovative firms that received public subsidies for R&D and related activities in 28 countries during the periods 2006-2008 and 2008-2010. Previous data from the same source shows that R&D subsidies are also important in the US (not included in Figure 1).

**Figure 1. Percentage of product and/or process innovative firms receiving public subsidies for innovation, 2006-08 and 2008-10 \***

Source: Adapted from OECD (2013, p.185)

\* Subsidies for innovation refer to ‘various financial incentives to encourage firms to engage in innovation activities (R&D and other)’ and exclude ‘indirect support (such as foregone revenue from R&D tax credits)’ OECD (2013, pp.184-185).

There is neither theoretically nor empirically definitive guidance on the effectiveness of public R&D subsidies in stimulating private R&D. Public support of private R&D decreases unit costs and increases the expected profitability of funded R&D projects thus giving an incentive for subsidised firms to invest in R&D activities over and above the counter-factual levels that firms would have undertaken without public support, leading to additionality. On the other hand, firms may substitute public funds for private funds that would have been committed in any case, leading to crowding out.

This paper provides a meta-regression analysis (MRA) of the quantitative microeconomic literature on the effectiveness of public R&D subsidies in triggering private R&D. Together, the primary studies, in their attempt to investigate the existence of a causal relationship between public R&D funding and private R&D, offer conflicting findings. Public subsidies are found to complement private R&D; to have no effect at all; or even to crowd-out private R&D. All three possible outcomes are well reported in the literature. In line with meta-regression studies of other literatures, we investigate this literature to determine:

1. the extent to which heterogeneous findings in this literature can be explained by the heterogeneity of samples and empirical methodologies;
2. the degree – if any – to which this literature is infected by publication selection bias; and
3. the genuine representative effect – if any – established by this literature after controlling for possible publication bias and sources of heterogeneity.

Previously, a meta-analysis (García-Quevedo, 2004) investigated the first of these concerns for studies published before or during 2002, concluding that no research dimensions, controlled for in the meta-analysis, can explain the heterogeneity in the empirical findings. The present study investigates studies published in or after 2000 and, in addition, investigates the second and third of these concerns. The estimation, controlling for publication bias, of a representative effect of public R&D subsidies on private R&D provides evidence on the effectiveness of public subsidies. Another distinguishing feature of this meta-regression analysis is the inclusion of studies not only on input additionality (i.e. additionality measured on R&D inputs like R&D expenditure and the number of employees committed to R&D activities) but also of studies on output additionality (i.e. additionality measured on R&D outputs like patents and new products).

# Theoretical context: competing perspectives

Different theories suggest contrasting outcomes. The mainstream perspective is that R&D has public good characteristics, namely non-rivalry and non-excludability.Consequently, the appropriability of private R&D outputs is not perfect, the corollary of which is lower private than social returns and a socially suboptimal level of R&D (Nelson, 1959; Arrow, 1962; Usher, 1964; Bloom et al., 2013). Moreover, due to inherent high risk in R&D investments, to the nature of R&D activities, which cannot be used as collateral in loan contracts, and to information asymmetry between creditors/investors and R&D performing firms, insufficient or too costly external capital is available to firms for financing R&D activities (Hall, 2002a,b). Both the public good characteristics of R&D and capital market imperfections constitute market failures, which provide the theoretical rationale for public intervention. Subsidies can mitigate these market failures by decreasing R&D unit costs and increasing the expected profitability of funded R&D projects, which incentivises greater private R&D spending; i.e. additionality. In other words, the effectiveness of subsidy in raising private R&D towards the socially optimum level precludes full crowding out. Yet, other theoretical perspectives suggest that intervention may lead to full crowding out: both self-interested selection procedures on the part of public bureaucracies and firms exploiting information asymmetries may reduce the effectiveness of subsidies.

From the perspective of public choice theory, public agencies may adopt opportunistic behaviour.[[2]](#footnote-2) In particular, it may be in their interests to support firms with R&D projects that are likely to succeed irrespective of public support. This “cherry-picking” strategy will result in apparent effectiveness, which will give credit to the agency’s managers, will justify the role of the agency itself and thus perpetuate its existence. Opportunistic bureaucrats may also support such a strategy in order to gain short-term reputation for the “effectiveness” of their programmes. Funding such projects, which are the ones most likely to be privately financed in the absence of public support, may also constitute a source of crowding out. Given that R&D subsidies may significantly increase the probability of a firm conducting R&D (Czarnitzki, 2006), cherry-picking may also contribute to the reinforcement of the already competent firms, without inducing new firms to undertake R&D.[[3]](#footnote-3)

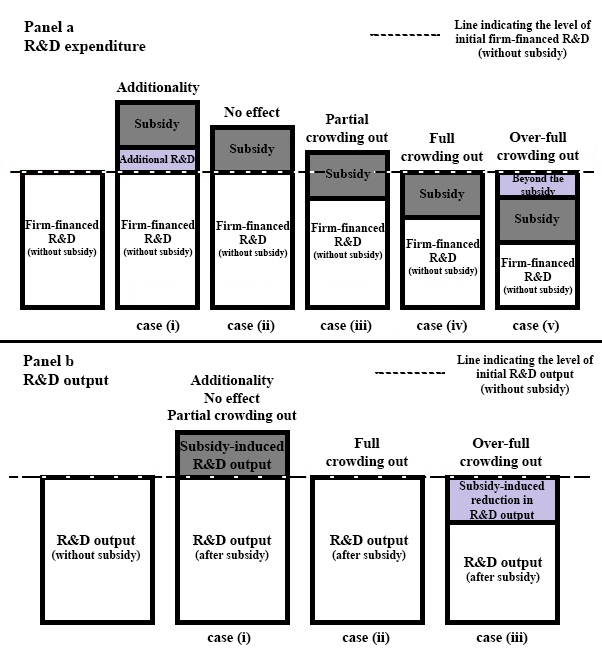
According to Aschhoff (2009), firms may view public funds as a relatively cheap way to finance their R&D projects, especially when application costs are low and the probability of selection is high compared to alternative financing sources. In addition, as is the case for many outputs of public policy (Butler, 2012, p.89), difficulties in measuring the private and social returns of R&D projects give rise to information asymmetry between public agencies and private firms. While Aschhoff (2009) points to the incentive for firms to hide private information from public agencies, asymmetric information enables firms to do so. Accordingly, we conjecture that, together, incentive and means render likely the suboptimal allocation of public funds. Between firms, hidden information potentially diverts public support to projects that would have proceeded in any case; while, within supported firms, hidden actions may include the diversion of allocated resources to uses other than those agreed.

In conclusion, theory alone cannot be conclusive regarding the size or even the direction of the effect of public R&D funding on private R&D. Hence, we turn to the corresponding empirical evidence. Yet, the extensive literature on the effects of public R&D subsidies on private R&D also yields ambiguous findings. Hence, we use MRA to analyse this literature as an ‘objective and critical methodology to integrate conflicting research findings’ (Stanley and Doucouliagos, 2012, p.2).[[4]](#footnote-4)

# Conceptual framework and heterogeneous empirical effects

In this section, we identify the different ways in which theoretically defined public subsidy effects relate to estimated outcome effects. This enables us to identify groups of studies within this heterogeneous literature that are sufficiently homogeneous for valid investigation by MRA. Figure 2 sets out the conventional conceptual framework for defining the range of possible outcomes of public support for both R&D expenditure (Panel a) and R&D output (Panel b).

**Figure 2. The possible effects of R&D subsidies on a) R&D expenditure and b) R&D output**



Source: own interpretation from the evaluation literature

Five possibilities for the effect of subsidy on R&D expenditure (input additionality) can be distinguished (Panel a):

1. *additionality* – the R&D subsidy triggers additional firm-financed R&D spending beyond the amount of the subsidy;
2. *no effect* – firms just add the amount of the subsidy to their firm-financed R&D expenditures, which remain unchanged;
3. *partial crowding out* – a part of the subsidy amount is used in place of own R&D expenditures, which become less than they would have been without the subsidy;
4. *full crowding out* – when firm-financed R&D expenditures decrease by the full amount of the subsidy; and
5. *over-full crowding out* – when firms decrease firm-financed R&D expenditures by more than the amount of the subsidy.

With respect to the effect of subsidy on R&D output (output additionality), only three possibilities can be distinguished (Panel b):

1. *additionality/no effect/partial crowding out* – the R&D subsidy triggers a higher level of R&D output than the counterfactual state of no subsidy but it cannot be determined whether additionality, no effect or partial crowding out takes place (because output measures such as patents and sales arising from innovation are not commensurable with the value of subsidies, estimated positive effects rule out full crowding out – Panel b, Case ii – but cannot distinguish between the three possible cases of subsidy-induced R&D output depicted in Panel b, Case i);
2. *full crowding out* – when R&D output remains at the same level as in the counterfactual state of no subsidy; and
3. *over-full crowding out* – when subsidised firms’ R&D output becomes smaller than in the counterfactual case.

Table 1 identifies each of the six combinations of outcome measure and subsidy measure to be found in the empirical literature and, for each combination, displays the possibility or otherwise of estimating the R&D subsidy effects detailed in Figure 2. For each combination, subsidy effects are estimated either as coefficients within a regression (parametric) framework or, equivalently, as average treatment on the treated effects (ATTs) within a matching (non-parametric) framework.

**Table 1. Possible effects identified depending on the type of outcome variable used and available information regarding the subsidy**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Outcome variable | Firm-financed R&D | | Total R&D | | Non-expenditure  Input R&D  (R&D employees) | R&D Output  (e.g. patents;  “R&D-sales”) |
|  | **Input Additionality** | | | | | **Output Additionality** |
|  | **1** | **2** | **3** | **4** | **5** | **6** |
| Information regarding amount of subsidy | **Value of subsidy known** | **Binary indicator only**  **(value of subsidy unknown)** | **Value of subsidy known** | **Binary indicator only**  **(value of subsidy unknown)** | **Value of subsidy known**  **/**  **Binary indicator only** | **Value of subsidy known**  **/**  **Binary indicator only** |
| The five possible effects of R&D subsidies \* |  |  |  |  |  |  |
| Additionality | ✓ | ✓ | ✓ | Not identifiable | Not identifiable | Not identifiable |
| No effect | ✓ | ✓ | ✓ | Not identifiable | Not identifiable | Not identifiable |
| Partial crowding out | ✓ | Not identifiable | ✓ | Not identifiable | Not identifiable | Not identifiable |
| Full crowding out | ✓ | Not identifiable | ✓ | ✓ | ✓ | ✓ |
| Over-full crowding out | ✓ | Not identifiable | ✓ | ✓ | ✓ | ✓ |
| Undifferentiated crowding out  (i.e. aggregating partial, full & over-full crowding out) | ✓ | ✓ | ✓ | n.a. | n.a. | n.a. |
| Number of studies  Reported estimates  % of total estimates | 12  115  12.48% | 15  185  20.09% | 4  15  1.62% | 24  345  37.46% | 5  22  2.40% | 19  239  25.95% |

Key: ✓- effect can be estimated or inferred; n.a. – not applicable  
\* The correspondence between Figure 2 and Table 1 is as follows:

* Panel a: Case i – Additionality; Case ii – No effect; Case iii – Partial Crowding Out; Case iv – Full Crowding Out; Case v – Over-full Crowding Out
* Panel b: Case i – Additionality, No effect and Partial Crowding Out jointly; Case ii - Full Crowding Out; Case iii - Over-full Crowding Out

Source: authors

Table 1 sets out the heterogeneity of this empirical literature, in which the effects reported by empirical studies reflect the different types of data available. For example, in order to identify all five cases for R&D expenditure (Figure 2a), researchers need either the value of firm-financed R&D and the value of the subsidy (which together gives the value of total R&D) or the value of total R&D and the value of the subsidy (the difference giving the value of firm-financed R&D). Only 15 studies (130 observations) meet these requirements (Table 1, Columns 1 and 3).[[5]](#footnote-5) However, the quantitative estimates from Column 1 and Column 3 studies cannot be pooled, because the effects sizes have different interpretations: for example, in Column 1, additionality is indicated by a significantly positive effect in either a regression or a matching context; but in Column 3 by an effect larger than the subsidy (matching) or by a subsidy coefficient larger than unity (regression).

Following Table 1, we partition the literature and our MRA database into three groups.

1. **Input I** (Columns 1 and 2). Around one-fifth of the total estimates reported in the literature use information on the value of firm-financed R&D but only binary information on the subsidy (i.e. whether or not subsidy has been received) (Table 1, Column 2). In this case, in both regression and matching contexts, a significant positive effect can be interpreted as additionality (Figure 2a, Case i); a non-significant effect as no effect (Figure 2a, Case ii); and a statistically significant negative effect as undifferentiated crowding out (i.e. a reduction in firm-financed R&D without being able to distinguish further among degrees or types of crowding out). We can pool the studies in Column 2 with those in Column 1, albeit with some loss of information from the latter.
2. **Input II** (Columns 3 and 4). The largest group of studies, accounting for almost 40 per cent of estimates in this literature, uses total R&D expenditure (i.e. including the amount of subsidy) as the outcome variable and a binary indicator for whether or not the firm receives a subsidy (Table 1, Column 4). We can pool the studies in Column 4 with those in Column 3, albeit with some loss of information from the latter. In this case, in regression and matching contexts alike, if the effect of subsidies on total R&D expenditure is estimated to be statistically significant and positive, then the only information that can be extracted is the exclusion of both full crowding-out and over-full crowding-out effects. A positive effect is equivalent to cases i, ii and iii jointly in Figure 2a, in all of which total R&D is greater than in the “without subsidy” – counterfactual – case; hence, we cannot separately identify partial crowding-out, no effect or additionality effects. In other words, this group follows the pattern depicted in Figure 2b, Case i. Conversely, when a statistically insignificant effect is estimated then full crowding out cannot be rejected (Figure 2a, Case iv). Lastly, if a statistically significant and negative effect is estimated, then over-full crowding out takes place (Figure 2a, Case v). Even though Column 5 follows a similar pattern to the pooled subsample of Column 3 and Column 4, its effect sizes are non-expenditure measures of R&D and so are not pooled with this subsample. Because there are too few estimates in this category for separate analysis, these are excluded from MRA.
3. **Output**. Just over a quarter of the estimates reported in this literature refer to the output effects (Column 6). This is a conceptually discrete subsample, since the effect sizes all refer to R&D outputs such as patents and sales arising from innovation. In this case, the subsidy amount does not improve on a binary indicator with respect to the identification of subsidy effects, because these are measured either in non-monetary terms (patents) or in monetary terms (sales) that bear no well-defined relationship with the value of the subsidy. Similar to Column 4, an insignificant estimated effect is evidence for full crowding out, a negative estimated effect evidence for over-full crowding out, and a positive effect precludes both full and over-full crowding out.

Table 2 displays counts of the effects identified in the three subsamples: in Input I, either separate effects or an aggregate crowding-out effect; and in Input II and Output either full and over-full crowding out separately or, conversely, an aggregate rejection of both full and over-full crowding out.[[6]](#footnote-6)

Table 2. Count of identified effects in the three MRA subsamples

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Input I** | | **Input II** | | **Output** | |
| **Effects** | *Separate* | *Aggregate* | *Separate* | *Aggregate* | *Separate* | *Aggregate* |
| **Additionality** | 164 |  | 11 | 312 |  | 152 |
| **No-effect** | 117 | 0 |
| **Partial**  **crowding out** | 5 | 6 | 4 |
| **Full**  **crowding out** | 0 | 33 |  | 79 |  |
| **Over-full**  **crowding out** | 8 | 0 | 8 |
| **Total** | 300 | | 360 | | 239 | |

According to Table 2, more than half of the estimates in the Input I subsample indicate R&D subsidy additionality effects (164); more than a third indicates no effect of R&D subsidy on firm-financed R&D expenditure (117); and only 19 from 300 indicate substitution of private expenditure by subsidy (crowding out of any degree). For the Input II subsample, there are no estimates indicating over-full crowding out while 33 suggest full crowding out. The largest part of this subsample reports positive estimates (327), of which 11 indicate additionality and four partial crowding out. However, most of these positive estimates (312) only preclude full or over-full crowding out. For the Output subsample, estimates pointing to full or over-full crowding out constitute almost two fifths of this group (87). Still, most estimates refer to positive effects (152) so that full and over-full crowding out can be rejected, although additionality, no effect and partial crowding out cannot be distinguished. To sum up, a first reading of the literature suggests that positive effects dominate and, in the cases where data is available, additionality seems to be achieved (Input I).

# MRA database and preliminary investigation

In this section, we discuss the criteria for study inclusion in the MRA database and the effect sizes to be used in the subsequent analysis.[[7]](#footnote-7)

## 4.1 Criteria for study inclusion

A previous meta-analysis conducted by García-Quevedo (2004) includes studies from 1966 to 2002. In contrast, the present MRA includes studies from 2000 to 2013. In 2000 the study of Busom (2000) was published, which is the first one in a series of studies that takes explicitly into account the endogeneity of the provision of public R&D funds. Thus, the year 2000 serves as a natural point of departure for our MRA database.

A search of the EconLit online database on the 11th of March 2013, using the keywords “R&D subsidies” and “R&D subsidies effect”, and refining the search to studies published in 2000 or later, returned 272 and 51 results respectively.Of course, not every one of those studies was relevant. From those studies, we coded only the ones in English and whose title and/or abstract explicitly stated that the effect of R&D subsidies on a measure of R&D was investigated. Further, only studies that use firm-level or plant-level data were included in the MRA database. Thereafter, references in some of these studies identified additional studies. Further, every study published in 2000 or after referenced by Zúñiga-Vicente et al. (2014), to our knowledge the most recent review of the literature, was also included in the MRA database.[[8]](#footnote-8) Only studies that involve direct R&D funding *inside* firms in the form of subsidies are included in the MRA database.

In total 52 studies investigating the effects of public R&D funding through subsidies (grants and/or loans) on a measure of R&D are included in the MRA database. (These studies are listed in Appendix A.) Since it is the norm for a study to report more than one estimate of interest, 921 effect sizes have been coded. Each study reports on average 18 effect sizes with the median being 12. The minimum number of effect sizes reported in a study is 1 whereas the maximum is 78.

## 4.2 Effect sizes (Partial Correlation Coefficients)

In the primary literature, there are four effect types to be coded: i) coefficients extracted from regressions (in level-level or log-level/semi-elasticities form); ii) marginal effects extracted from non-linear models (negative binomial regressions, probit and tobit models); iii) constant elasticities (log-log specification); and iv) average treatment on the treated (ATT) effects (mean difference of treated and control groups).

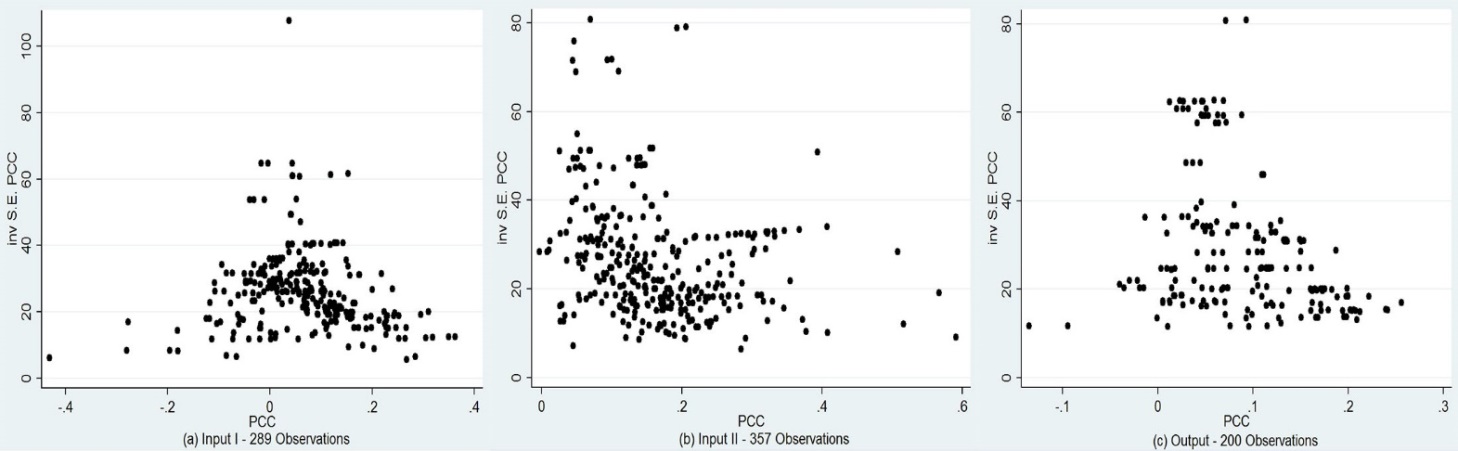
This heterogeneity in reported effects has previously been noted, for example, by David et al. (2000). To make these effects comparable so as to be able to combine them, we follow Doucouliagos and Stanley (2009) and Stanley and Doucouliagos (2012) by transforming these effects into partial correlation coefficients (PCCs). The PCC is a unit-free measure of the magnitude and direction of the association between two variables (public R&D subsidies and a measure of R&D outcomes) holding other variables constant. However, given that the PCC is applicable to a linear association between two variables, we exclude 43 effect sizes derived from non-linear models from the subsequent analysis.[[9]](#footnote-9) Its calculation, along with that of its standard error, is straightforward:

(1) (2)

where *t* stands for the t-statistic on the estimated subsidy effect and *df* for the degrees of freedom extracted from the respective estimate in the primary literature.

Funnel plots for all three subsamples are presented in Figure 3 below.[[10]](#footnote-10) These plot the effect size (PCC) from each estimate reported in the literature against its precision (the inverse of the standard error). If the literature is investigating a relationship with some true or at least representative effect, then from sampling theory we expect large sample studies to yield relatively precise estimates tightly distributed around the true effect and small sample studies to yield less precise estimates widely dispersed around the true effect. In this case, the plot is symmetric around the true effect, resembling an inverted funnel. Conversely, a skewed distribution can reflect efforts (e.g. specification search) to obtain estimates of a certain sign or – especially in small sample studies – to obtain larger estimates to offset larger standard errors. Asymmetries in the funnel plot are thus prima facie evidence that the literature is affected by publication bias. Estimating and then controlling for publication bias to identify representative effect sizes “beyond publication bias” in empirical literatures is the main contribution of meta-regression analysis to the analysis of quantitative literatures (Stanley, 2005).

**Figure 3. Funnel plots for each subsample: a) Input I; b) Input II; and c) Output**



The funnel plot for the Input I subsample (Panel a) seems to be more or less symmetric around a small positive PCC, which indicates a small average additionality effect. In contrast, the plot for the Input II subsample (Panel b) reflects the different character of this subsample, in which statistically significant and positive reported effects merely preclude full and over-full crowding out. In this case, the most precise PCCs are mainly clustered around a value of slightly less than 0.10, while the distribution of PCC values around this spout is highly skewed to the right, suggesting the presence of pronounced positive publication bias.[[11]](#footnote-11) Inspection of funnel plots can indicate sources of heterogeneity that require further investigation. For example, in this subsample (Panel b) visual inspection reveals two precisely estimated PCCs (inverse SE of c.80) with values of around 0.20, whereas the main cluster of the most precisely estimated PCCs have estimated PCCs of around 0.10. Both of these are derived from regressions that do not take account of the potential endogeneity of R&D subsidies and so may be substantial overestimates of the true effect. We investigate this possibility further in the context of both bivariate and multivariate MRA below. Finally, Panel (c) for the Output subsample shows the main body of the distribution centred around a small positive PCC with, at most, a slight right skew, suggesting that R&D subsidies do not lead to full or over-full crowding out, even though this part of the literature may be affected by positive publication bias (i.e. over-reporting of estimates excluding full- and over-full crowding out).

In Table 3, both unweighted and weighted averages of the PCCs for each group are reported (in Columns 1 and 3 respectively). In Column 3, each PCC is weighted by the inverse of its variance, which has been shown to be the optimal weighting scheme (Cooper and Hedges, 1994). Thus, more precise estimates are given greater weight, since these estimates are expected to more closely approximate the respective population parameters.

**Table 3. Unweighted and weighted averages of PCCs by subsample**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Average**  **Subsample** | **Unweighted** | **Unweighted**  **TOP-10%** | **Weighted**  **(inverse**  **variance)** | **Weighted**  **(inverse**  **variance)**  **TOP-10%** |
|  | 1 | 2 | 3 | 4 |
| **Input I** | 0.065  [.053-.077]  289  observations | 0.061  [.041-.080]  29 observations | 0.054  [.046-.063]  289 observations | 0.052  [.033-.071]  29  observations |
| **Input II** | 0.158  [.149-.168]  357  observations | 0.099  [.076-.123]  36 observations | 0.131  [.122-.140]  357 observations | 0.100  [.077-.123]  36  observations |
| **Output** | 0.089  [.080-.099]  200  observations | 0.049  [.039-.059]  20 observations | 0.070  [.063-.076]  200 observations | 0.051  [.040-.062]  20  observations |

Notes. 95% Confidence Intervals are reported within brackets

The effects located at the top of the funnel plots are the most precise and the least affected by publication selection bias, because greater precision reduces the incentive for researchers to search across datasets, model specifications and methodologies for “significant” results. Accordingly, simulations have shown that the average of the 10 per cent most precise effect sizes provides a statistic that is by and large not distorted by publication selection bias (Stanley et al., 2010). In Table 3, Columns 2 and 4 respectively report the unweighted and weighted averages of the 10 per cent most precise PCCs for each subsample. In each subsample, the average of the 10 per cent most precise effect sizes points to a smaller effect of R&D subsidies than when every estimate is taken into account. This is suggestive of positive publication selection bias.

The averages in Table 3 are consistent with the funnel plots; in each subsample, a positive but small to moderate PCC and an indication of positive publication bias.[[12]](#footnote-12) The average PCCs suggest that Input I estimates typically support additionality, while both Input II and Output estimates reject full and over-full crowding out. However, no safe conclusions can be reached by using only information from these unconditional averages. Potential publication selection bias and systematic heterogeneity may be present, thus distorting the average effect size. Accordingly, we move beyond this preliminary investigation first with bivariate MRA (Section 5) and then with multivariate MRA (Section 6).

# Bivariate MRA: genuine empirical effect beyond publication selection bias

## 5.1 FAT – PET – PEESE

In this section, we follow Stanley (2008) and Stanley and Doucouliagos (2012) to identify publication selection and to estimate a genuine empirical effect beyond publication selection bias. When publication selection is absent, the effect sizes are independent from their standard errors (Stanley and Doucouliagos, 2012), which implies the following equation:

(3)

where *i*=1,…, n indexes the n individual estimates reported in the primary literature. *SEi* denotes the standard error of the *i-*th *PCC* and *ε* is the usual regression error. In the presence of publication selection, *β1* will be statistically significant and its sign gives the direction of publication bias. Moreover, as standard errors converge to zero (i.e. precision asymptotes towards infinity) the expected value of the effect sizes converges to *β0*. Thus, testing for *β1*=0 constitutes a test for the presence of publication selection (the so called Funnel Asymmetry Test – FAT) and testing for *β0*=0 for the existence of a genuine effect (the so called Precision Effect Test – PET). In the case that the null hypothesis, H0: *β0*=0 – i.e. no genuine effect – is rejected, then the magnitude of *β0* is an estimate for the genuine effect beyond publication selection bias. However, Equation (3) suffers from heteroskedasticity; the variance of *PCCi* (and thus *εi*) is not typically constant. To correct for heteroskedasticity, Weighted Least Squares (WLS) estimation is employed by weighting the squared errors with the inverse variance of each effect size or, equivalently, by estimating Equation (4) below, which is Equation (3) divided by the standard error of *PCCi*(*SEi*):

(4)

In Equation (4), *ti* is the t-statistic of each PCC (which is the one extracted from the primary literature) and *νi=εi/SEi* is the new error term corrected for heteroskedasticity.[[13]](#footnote-13) However, the coefficient on precision, *β0*, in Equation (4) is biased in the presence of publication selection and simulations have shown that in the presence of an authentic empirical effect it is better to use the variance instead of the standard error in Equation (3) (Stanley and Doucouliagos, 2007, 2011; Moreno et al., 2009). Thus, Equation (5) below, which also takes account of heteroskedasticity, provides a better estimate of the underlying effect, when there is one, corrected for publication selection bias:

(5)

Testing the null hypothesis H0: *β0*=0 in Equation (5) provides the so called Precision Effect Estimate with Standard Error (PEESE) test.

Studies typically report more than one effect size. To explicitly take account of otherwise unmodelled between-study heterogeneity, we estimate Equation (4) as a fixed effect (FE) unbalanced panel model:

(6)

where *tis,* *SEis* and *uis* denote the t-statistic, standard error and residual on the *i*-th effect in the *s*-th study, respectively, and *μs* denotes the study-specific fixed effects. There is some evidence that FE estimation yields somewhat biased estimates compared to the WLS model (Equation (4)) (Koetse et al., 2010; Doucouliagos and Stanley, 2013). However, debate continues on this issue and the current consensus is that best practice in MRA is to estimate and report alternative models (Stanley and Doucouliagos, 2012, p.104; Stanley et al., 2013). Accordingly, we employ fixed-effects estimation as a robustness check on WLS estimation, which is the most common approach to MRA.[[14]](#footnote-14)

Table 4 reports estimates for each subsample – Input I, Input II and Output – from Equation (4) (the FAT and the PET), Equation (5) (PEESE version of PET only) and Equation (6) (the FAT and the PET). In each case, cluster-robust standard errors are used, which allow for potential dependence among effect sizes from the same study; this approach constitutes a conservative approach to inference (higher standard errors) on the MRA coefficients.

The Input I estimates are consistent with the corresponding funnel plot (Figure 3a). When fixed effects are controlled for (Equation (6)), the FAT yields evidence of positive publication bias similar to the FAT estimate for Equation (4) in magnitude but significant at the 5 per cent level. Similarly, the PET does not reject the existence of a genuine effect under Equation (6) but it does under Equation (4), although the PCC estimates are almost identical; in both cases, according to the criteria suggested by Doucouliagos (2011) (see Footnote 12), this effect is “small”. The PEESE is not taken into account, since the PET under Equation (4) does not identify a statistically significant effect. The statistical significance of one estimate and the similar size of the other suggest additionality from R&D subsidy, although this effect is small. Moreover, both of these conditional PET estimates (0.026 and 0.028) are smaller than the unweighted average of 0.065 (Table 3), which is consistent with positive publication bias.

Consistent with the corresponding funnel plot (Figure 3b), in the Input II subsample the FAT-PET-PEESE procedures provide uniform evidence of both positive publication bias and positive genuine effect of “moderate” size beyond publication bias.[[15]](#footnote-15) Although the deviation of all three PET/PEESE estimates (0.075, 0.110 and 0.101) from the unweighted average of 0.158 (Table 3) is again consistent with positive publication bias, each estimate of the PET clearly suggests the exclusion of full and over-full crowding out.

In the Output subsample, a positive genuine effect of “small” size is suggested by both the WLS PET-PEESE procedure and by the fixed effects PET estimate. The impression of the funnel plot suggesting positive publication bias in this subsample (Figure 3c) is supported by the FAT test resulting from Equation (4), although the estimated publication bias loses statistical significance once fixed effects are controlled for. However, we are inclined to accept the evidence of publication bias, given that the deviation of all three PET/PEESE estimates (0.035, 0.055 and 0.054) from the unweighted average of 0.089 (Table 3) is consistent with positive publication bias.[[16]](#footnote-16)

**Table 4. FAT, PET and PEESE tests for Equations (4), (5) and (6)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model**  **Subsample** | **FAT (β1=0)**  **WLS Equation (4)** | **PET (β0=0)**  **WLS**  **Equation (4)** | **PEESE** a **(β0=0) WLS**  **Equation (5)** | **FAT(β1=0)**  **FE estimation**  **Equation (6)** | **PET(β0=0)**  **FE estimation**  **Equation (6)** |
| **Input I**  289 observations | 0.89  [-.74, 2.52]  (t=1.13) | 0.026  [-.024, .077]  (t=1.08) | 0.046\*\*\*  [.016, .075]  (t=3.24) | 0.85\*\*  [.11, 1.58]  (t=2.39) | 0.028\*\*  [.00009, .0557]  (t=2.08) |
| **Input II**  357 observations | 1.87\*\*\*  [1.16, 2.58]  (t=5.42) | 0.075\*\*\*  [.043, .107]  (t=4.77) | 0.110\*\*\*  [.082, .138] (t=8.11) | 1.17\*\*  [.047, 2.30]  (t=2.14) | 0.101\*\*\*  [.059, .143]  (t=4.95) |
| **Output**  200 observation**s** | 1.28\*\*  [.25, 2.31]  (t=2.65) | 0.035\*\*  [.007, .063]  (t=2.68) | 0.055\*\*\*  [.036-.075] (t=6.08) | 0.74  [-.51, 1.98]  (t=1.26) | 0.054\*\*  [.011, .097]  (t=2.66) |

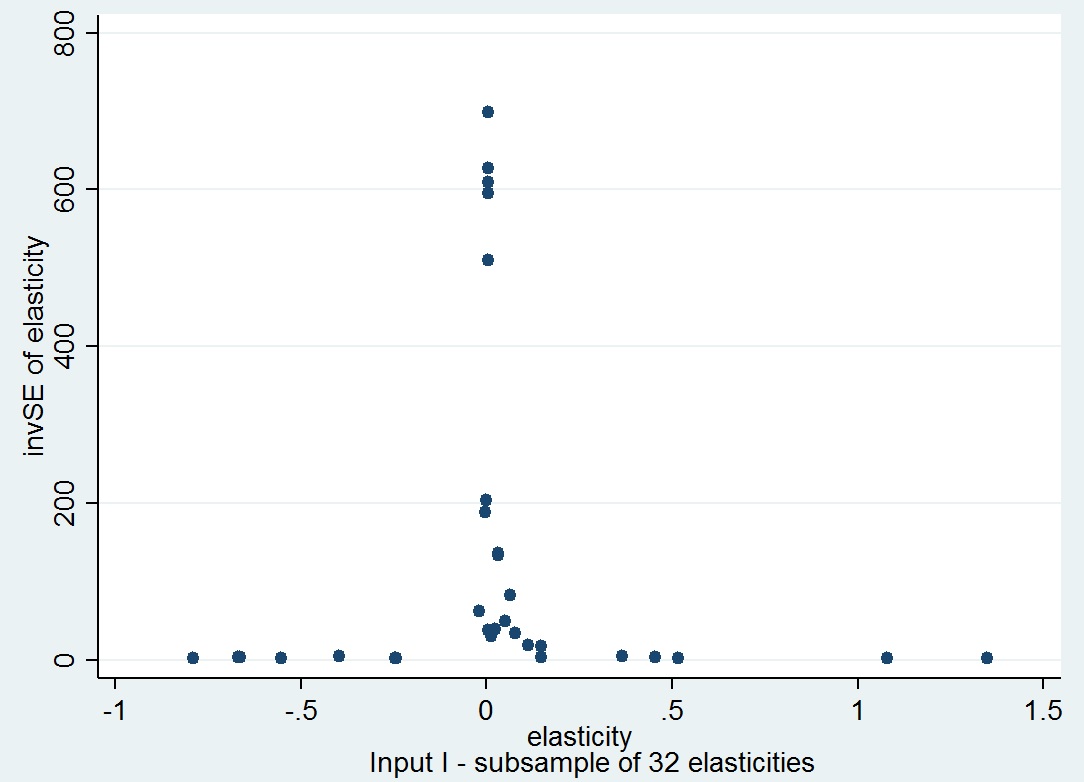
Notes. t-values reported in parentheses are calculated from cluster-robust standard errors. 95% Confidence Intervals are reported in brackets. \*\*\* and \*\* indicate statistical significance at the 1% and 5% levels respectively

a PEESE should be interpreted only when the PET yields a significant result.

In sum, according to the guidelines of Doucouliagos and Stanley (2013), bivariate MRA suggests “little to modest” publication bias (FAT<1) in the Input I subsample, “substantial” publication bias (1≤FAT≤2) in the Input II subsample, and mixed evidence in the Output subsample. In each case, the evidence of positive publication bias from these conditional estimates is consistent with the evidence from the unconditional PCC averages reported in Table 3. With respect to the PET/PEESE estimates, we find evidence of “small” to “moderate” effects of R&D subsidies (see Footnote 12), which provide evidence of either additionality (Input I subsample) or an absence of crowding out (Input II and Output subsamples).

## 5.2 Elasticities as effect sizes

To gain further insight into the size of R&D subsidy effects, we analyse a subsample of Input I for which 32 constant elasticity measures from 6 studies are available.[[17]](#footnote-17) Elasticities – the percentage response of firms’ R&D spending to a percentage change in subsidy – from the primary studies yield comparable effect sizes without transformation. These elasticities are plotted against their inverse standard errors in Figure 4.

**Figure 4. Funnel plot of 32 elasticities against their inverse standard errors** 

The main body of the elasticities is centered around zero while the distribution seems slightly skewed to the right. The above funnel plot indications are empirically confirmed by the corresponding FAT-PET and PEESE implemented by estimating Equations (7), (8) and (9), which are analogous to Equations (4), (5) and (6):

(7)

(8)

(9)

where *ei*stands for the *ith* elasticity; *eis* for the *ith* elasticity in the *sth* study; is the standard error of the *ith* elasticity and is the standard error of the *ith* elasticity in the *sth* study; and *μs* stands for the study specific fixed effects. In these models, the estimated effect size directly gives the elasticity representative of this subsample. Table 5 reports the estimates.

**Table 5. FAT, PET and PEESE tests for Equations (7), (8) and (9)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Elasticity**  **Subsample** | **FAT**  **(β1=0)**  **WLS**  **Equation (7)** | **PET**  **(β0=0)**  **WLS**  **Equation (7)** | **PEESE**  **(β0=0)**  **WLS**  **Equation (8)** | **FAT**  **(β1=0)**  **FE estimation**  **Equation (9)** | **PET**  **(β0=0)**  **FE estimation**  **Equation (9)** |
| **Input I**  32 observations | 0.727\*  [-0.117, 1.571]  (t=1.76) | 0.0034\*\*  [0.0003, 0.0064]  (t=2.27) | 0.0048\*\*\*  [0.0021, 0.0074]  (t=3.65) | 0.063  [-1.070, 1.196]  (t=0.11) | 0.0078\*\*  [.0015, .0141]  (t=2.53) |

Notes. t-values reported in parentheses are calculated from default standard errors. The conservative approach to inference demands reporting default SEs, since in this small subsample cluster-robust and heteroskedasticity-robust SEs are considerably smaller than the default SEs. 95% Confidence Intervals in square brackets.

\*\*\*, \*\* and \* denote statistical significance at the 1%, 5% and 10% levels respectively.

Compared to the full-sample estimates for Input I reported in Table 4, these subsample estimates of the FAT provide weaker evidence of publication bias: in Table 4, although the WLS estimate (0.89) is not statistically significant, it is similar to the significant FE estimate (0.85; t=2.39); in Table 5, the WLS estimate (0.727) is significant only at the 10 per cent level and is very different from the FE estimate, which cannot be distinguished from zero (0.063; t=0.11). Conversely, whereas the full Input I subsample does not yield uniform support for an authentic positive effect (only the FE estimate of the PET is statistically significant, although the WLS estimate is of similar size), in this subsample all three PET/PEESE estimates provide evidence of a statistically significant and positive empirical effect of R&D subsidies on input R&D.[[18]](#footnote-18) However, these effects are rather too small to be economically significant, representing estimated elasticities of less than 0.01; according to the PEESE (Equation 8) and PET (Equation 9) estimates, a 100 per cent increase in R&D subsidies (to illustrate using a non-marginal effect) causes, respectively, a 0.48 per cent and a 0.78 per cent increase in the private R&D expenditure of the subsidised firm. Nonetheless, although small, this increase in private R&D is evidence of additionality (see Figure 2a, Case i – “Additional R&D”). These estimated effects are in line with the rather small PCCs reported for the Input I subsample in Table 4 above (see Footnote 12).

# Multiple MRA

## 6.1 Heterogeneity

In this section, we employ multivariate MRA models to investigate sources of effect size heterogeneity. The excess heterogeneity of reported effect sizes (i.e. heterogeneity that is not due to the sampling error alone) is clearly revealed by Cochran’s Q-statistic (p=0.000). We now model this observed heterogeneity by relating it to heterogeneities in the primary studies with respect to the data used, methodologies employed as well as in contextual features of the research process. These hypothesised sources of this excess heterogeneity are incorporated into the specification of the bivariate MRA models estimated so far, as “moderator variables”, in order to obtain a better understanding of the variation of the effect sizes (PCCs). However, only those research dimensions present in at least three primary studies are specified in the MRA models.

Sources of heterogeneity in the samples used by the primary studies are captured by the following moderator variables: panel data, which enables unobservable influences at firm level to be controlled and, thus, sources of potential endogeneity to be addressed (although such data is seldom available to researchers in this literature; see Appendix A, the compendium table); data from either the Community Innovation Survey or the Mannheim Innovation Panel (CIS/MIP; respectively, a large-sample cross-section survey and an associated panel); firm-level as opposed to plant-level data; developed economies; European Union economies; the mid-point of the data (allowing for changes in either real underlying effects over time and/or for changes in research practices); data span (greater span reduces bias from time-specific shocks); data consisting of micro and SMEs only; manufacturing sector only; high-tech only; R&D performers only; innovators only; subsidies from central governments only; and binary indicator for the receipt of subsidy (binary treatment). In addition, we investigate potential heterogeneity in estimated effect sizes arising from the following aspects of research design and method: propensity score matching (PSM), which controls for observable differences between firms but is not able to address unobservable heterogeneity; difference-in-differences estimation (allowing for unobservable heterogeneity to be controlled); instrumental variable estimation (the conventional approach to endogeneity issues); subsidies in a previous period (a common way to address potential endogeneity in applied research); log-log specification; whether or not an estimate is reported as a main result or as a robustness check; methods not controlling for the potential endogeneity of subsidies (mainly employed for comparison purposes and not frequently used in the current literature; see Table 6 below); and year of publication (accounting for methodological innovation, improved data sets and other possible sources of changing research quality over time). Finally, we include two moderator variables that may capture possible contextual features of the research process – whether or not a study is published in a peer-reviewed journal or by the ZEW Research Centre – together with one that is of purely contextual interest, namely, whether or not a study acknowledges financial support (to check for possible systematic sources of non-scientific influence on the results reported in the literature). Precise definitions together with standard descriptive statistics for each moderator variable across all three subsamples are given in Table 6 below; moreover, some of these moderator variables are further discussed when interpreting the multivariate results reported in Table 7 below.

In the MRA literature, it is common practice to distinguish between moderator variables capturing heterogeneous sample and methodological influences on estimated effects and moderators capturing contextual influences reflected in publication bias. We follow Doucouliagos and Stanley (2009) in denoting the former as *Z*-variables and the latter as *K*-variables. Both model (4) and model (6) can be expanded to include variables that potentially moderate the estimated effect of R&D subsidies on measures of R&D:

(10)

(11)

where , and , stand for the *m* *K*-variables and *n* *Z*-variables that model sources of publication selection and heterogeneity of the effect sizes (PCCs), respectively, with and being their respective coefficients. Acknowledging that this distinction is somewhat arbitrary (Stanley and Doucouliagos, 2012, p.91), all but one of our moderator variables potentially explain heterogeneity of the effect sizes (i.e. *Z*-variables, which interact with the inverse of the standard error of the PCC) whereas only a few are assumed also to reflect aspects of the research and publication process that might contribute to publication bias (*K*-variables, which interact with the intercept).

**Table 6. *Z* and *K* moderator variables with descriptions, means and standard deviations for each subsample**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| No. | Moderator variables | Description | *Z or K* | Input I  mean  (standard deviation) | Input II  mean  (standard deviation) | Output  mean  (standard deviation) |
| 1 | **Year of Publication** | = a continuous moderator variable within the interval 2000 to 2013 | *Z* | 2007.5  (3.2) | 2008.4  (3.0) | 2008.1 (2.6) |
| 2 | **Published Study** | =1 when study published in peer-reviewed journals, 0 otherwise | *Z and K* | .7059 (.4564) | .6162  (.4870) | .2900 (.4549) |
| 3 | **ZEW Research Centre** | =1 if study comes from ZEW, 0 otherwise | *Z and K* | .1592 (.3665) | .4482  (.4980) | .5600 (.4976) |
| 4 | **Financial Support** | =1 if authors acknowledge financial support, 0 otherwise | *K* | .5502 (.4983) | .7871  (.4099) | .5900 (.4931) |
| 5 | **Panel Data** | =1 if study uses panel data, 0 if uses cross-section data | *Z and K* | .2734 (.4465) | .0980  (.2978) | .1000 (.3008) |
| 6 | **CIS/MIP Data** | =1 if study uses data from CIS or MIP, 0 otherwise | *Z* | .2007 (.4012) | .7199 (.4497) | .3450 (.4766) |
| 7 | **Firm Level**  (no variation in Input II) | =1 if study uses firm-level data, 0 if uses plant-level data | *Z* | .8478 (.3599) | 1  (0) | .6400 (.4812) |
| 8 | **Developed Economies** | =1 if effect sizes come from developed economies, 0 otherwise | *Z* | .8858 (.3186) | .8459 (.3615) | .6850 (.4657) |
| 9 | **European Union** | =1 if effect sizes come from economies within the EU, 0 otherwise | *Z* | .7855 (.4112) | .9580 (.2009) | .9200 (.2720) |
| 10 | **Mid-point of data** | = a continuous moderator variable within the interval 1986 to 2005 | *Z* | 1998 (4.18) | 1999.6 (2.67) | 2000.6 (2.06) |
| 11 | **Data Span** | = a continuous moderator variable within the interval 1 to 14 years | *Z* | 5.83  (3.60) | 4.76  (2.86) | 4.75  (4.15) |
| 12 | **Micro & SMEs** | =1 when micro, small or medium firms only are included in the sample, 0 when large firms are present in the sample | *Z* | .1661 (.3728) | .1849 (.3887) | .2800 (.4501) |
| 13 | **Manufacturing Sector only** | =1 if firms come from manufacturing sector only, 0 if from services or from both manufacturing and services | *Z* | .5329 (4998) | .1625 (.3694) | .1150 (.3198) |
| 14 | **High-Tech only**  (not in Output) | =1 if only high or medium-high technology firms are included in sample, 0 if i) low/medium-low tech or ii) high/medium-high and low/medium-low sectors are both present in the sample. | *Z* | .2180 (.4136) | .1625 (.3694) | 0  (–) |
| 15 | **R&D performers only** | =1 if only R&D performers are included in the sample, 0 otherwise | *Z* | .4810 (.5005) | .2101 (.4079) | .1550 (.3628) |
| 16 | **Innovators only** | =1 if only innovators are included in the sample, 0 otherwise | *Z* | .0450 (.2076) | .3866 (.4876) | .2250 (.4186) |
| 17 | **PSM method** | =1 if PSM method is used, 0 otherwise | *Z and K* | .5294  (.5) | .6303 (.4834) | .7350 (.4424) |
| 18 | **Difference-in-Differences**  (not in Output) | =1 if DiDmethod is used, 0 otherwise a | *Z and K* | .1730 (.3789) | .0728 (.2602) | 0  (–) |
| 19 | **Instrumental Variable**  (not in Output) | =1 if IV approach is followed, 0 otherwiseb | *Z and K* | .1280 (.3347) | .0280 (.1652) | 0  (–) |
| 20 | **Subsidies from Central Governments only** | =1 if subsidies come from Central Governments only, 0 otherwise | *Z* | .3702 (.4837) | .1120 (.3159) | .0800 (.2720) |
| 21 | **Subsidies in previous period** | =1 if subsidies provided in a previous period than R&D or innovation is measured, 0 if contemporaneous | *Z* | .2111 (.4088) | .0308 (.1731) | .4450 (.4982) |
| 22 | **Binary treatment** | =1 if subsidies are measured by binary variables, 0 otherwise | *Z* | .6401 (.4808) | .9608 (.1944) | .8350 (.3721) |
| 23 | **Log-log specification** (not in Input II and Output) | =1 if dependent variable and subsidies are measured in logs, 0 otherwise | *Z and K* | .1073 (.3100) | 0  (–) | 0  (–) |
| 24 | **Robustness check** | =1 if estimates are robustness checks, 0 if better described as main results | *Z and K* | .2872 (.4532) | .3137 (.4647) | .2750 (.4476) |
| 25 | **No-control of endogeneity**  (not in Output) | =1 if primary estimates come from models not controlling for endogeneity, 0 otherwise | *Z and K* | .0415 (.1998) | .1064 (.3088) | 0  (–) |

Notes.

a The moderator variable DiD refers to difference-in-differences with matching and difference-in-differences alone.  
b The moderator variable IV refers to selection models using IV estimation and IV estimation alone.

Table 6 highlights an important difference between MRA and conventional, narrative literature review. Whereas multiple MRA takes the whole literature into account to relate the heterogeneity of samples and empirical methodologies to heterogeneous findings, narrative literature review approaches the same task by careful comparison and contrast between selected studies. However, this demarcation need not be rigid. To bring the reader closer to individual studies within the literature – hence, to better convey the sense of the heterogeneity between particular studies or small groups of studies – we provide a compendium table of the sort typically provided in narrative reviews (Appendix A). Narrative review also enables the author to highlight differences between studies that are deemed to be important but that are associated with too few studies to be investigated by MRA. For example, while most studies treat R&D as inseparable due to data availability (Czarnitzki et al., 2011), Clausen (2009) and Czarnitzki et al. (2011) both find that subsidy of “far from the market” research yields a more positive impact than does subsidy of “close to the market” development activities. Similarly, Ozcelik and Taymaz (2008) investigate the effectiveness of subsidies both in the form of grants and in the form of loans, finding in both cases that full crowding out is excluded.

## 6.2 Multiple MRA estimation

The estimation results after estimating Equations (10) and (11) are summarised in Table 7. Both equations are additionally weighted by the inverse number of effect sizes reported in each study, so that no study reporting more estimates than others exerts a greater influence on the MRA. Cluster-robust standard errors are used with studies being the clusters. Initially, all *Z*- and *K*-variables are included in each specification. For each specification separately, the model is estimated and the least significant moderator is dropped from the model; the process is iterated until no statistically insignificant moderators (at the 10% confidence level) remain in the specification.[[19]](#footnote-19)

**Table 7. Multiple MRA estimation results for all three subsamples**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Input I | | Input II | | Output | |
|  | WLS | FE | WLS | FE | WLS | FE |
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) |
| **Z-variables** | | | | | | |
| *invsepcc* | -16.64\*\*\* | -19.87\*\* | 0.0143 | 0.283\*\*\* | -16.07\*\*\* | -11.08 |
| (inverse SE of the PCC) | (2.754) | (7.234) | (0.0572) | (0.0469) | (2.354) | (8.060) |
| *invSEyearofpublication* | 0.00832\*\*\* | 0.00576\* |  |  |  | -0.0182\*\*\* |
| (year of publication) | (0.00137) | (0.00303) |  |  |  | (0.00229) |
| *invSEpanel* |  | 0.318\*\*\* | 0.121\*\*\* | 0.0972\*\*\* | -0.124\*\*\* | 0.194\*\* |
| (panel data) |  | (0.0410) | (0.0411) | (0.0311) | (0.00993) | (0.0798) |
| *invSEfirmlevel* |  | 0.121\*\*\* |  |  | -0.0443\*\*\* |  |
| (firm-level data) |  | (0.0281) |  |  | (0.00821) |  |
| *invSEmid*\_*point* |  | 0.00412\*\*\* |  |  | 0.00801\*\*\* | 0.0237\*\*\* |
| (mid-point of data) |  | (0.00116) |  |  | (0.00117) | (0.00379) |
| *invSEdataspanyears* | 0.00353\*\* | -0.00746\*\* |  | -0.00555\*\*\* |  |  |
| (data span) | (0.00148) | (0.00346) |  | (0.00149) |  |  |
| *invSEmicro*\_*smes* |  | -0.0340\*\*\* |  |  |  |  |
| (micro & SMEs) |  | (0.00980) |  |  |  |  |
| *invSEhigh*\_*tech* |  | -0.0223\*\* |  |  |  |  |
| (high-tech sector) |  | (0.00804) |  |  |  |  |
| *invSErdperformersonly* |  | -0.0160\* |  | -0.0201\*\* |  | 0.0594\*\*\* |
| (R&D performers only) |  | (0.00779) |  | (0.00977) |  | (0.00544) |
| *invSEpsm* | -0.0266\*\* | 0.0268\*\* | 0.0501\* |  |  | 0.338\*\*\* |
| (PSM method) | (0.0114) | (0.00953) | (0.0286) |  |  | (0.0874) |
| *invSEdid* |  | -0.217\*\*\* | 0.0521\*\* | -0.0288\* |  |  |
| (D-i-D method) |  | (0.0291) | (0.0202) | (0.0144) |  |  |
| *invSEno\_control\_endogeneity* | 0.0359\*\*\* | 0.0729\*\*\* | 0.227\*\*\* | 0.0671\*\*\* |  |  |
| (not addressing endogeneity) | (0.0123) | (0.0137) | (0.0366) | (0.0156) |  |  |
| *invSEdeveloped* | -0.0815\*\*\* |  |  |  |  |  |
| (developed economy) | (0.0167) |  |  |  |  |  |
| *invSEeu* | 0.0294\* |  | 0.127\*\* |  | 0.139\*\*\* |  |
| (EU economy) | (0.0167) |  | (0.0458) |  | (0.0140) |  |
| *invSEiv* | 0.0778\*\*\* |  |  | -0.0690\*\*\* |  |  |
| (IV method) | (0.0230) |  |  | (0.0217) |  |  |
| *invSElog\_log* | -0.0947\*\*\* |  |  |  |  |  |
| (log-log specification) | (0.0147) |  |  |  |  |  |
| *invSEzewpaper* |  |  | -0.0659\*\* |  |  |  |
| (ZEW study) |  |  | (0.0283) |  |  |  |
| *invSEmanufacturing* |  |  | 0.0450\*\* |  | 0.0390\*\*\* |  |
| (manufacturing sector) |  |  | (0.0188) |  | (0.00677) |  |
| *invSEprevious\_period* |  |  | -0.114\*\* |  |  |  |
| (subsidy in previous period) |  |  | (0.0436) |  |  |  |
| *invSEbinary* |  |  | -0.104\*\* | -0.222\*\*\* |  | -0.00744\*\*\* |
| (binary measurement of subsidy) |  |  | (0.0467) | (0.0418) |  | (0.00157) |
| *invSEcismip* |  |  |  | 0.0337\*\* |  |  |
| (data from CIS/MIP) |  |  |  | (0.0154) |  |  |
| *invSErobustness\_check* |  |  |  | 0.0884\*\*\* | 0.0314\*\*\* |  |
| (robustness checking) |  |  |  | (0.0239) | (0.00917) |  |
| *invSEpublished* |  |  |  |  | 0.0491\*\*\* | 0.0780\*\*\* |
| (published study) |  |  |  |  | (0.00880) | (0.0176) |
| **K-variables** | | | | | | |
| *panel* |  | -12.80\*\*\* | -2.672\*\*\* |  |  | -12.75\*\*\* |
| (panel data) |  | (2.213) | (0.756) |  |  | (3.159) |
| *did* | -1.456\* | 8.879\*\*\* |  |  |  |  |
| (D-i-D method) | (0.751) | (1.137) |  |  |  |  |
| *iv* | -3.098\*\*\* | -1.098\*\*\* |  |  |  |  |
| (IV method) | (0.759) | (0.230) |  |  |  |  |
| *log\_log* | 2.832\*\*\* | -1.945\*\*\* |  |  |  |  |
| (log-log specification) | (0.594) | (0.0935) |  |  |  |  |
| *no\_control\_endogeneity* |  | -1.913\*\*\* | -4.007\*\*\* |  |  |  |
| (not addressing endogeneity) |  | (0.678) | (1.178) |  |  |  |
| *zewpaper* | 1.316\*\*\* |  | 1.031\* |  |  |  |
| (ZEW study) | (0.456) |  | (0.551) |  |  |  |
| *financialsupportstatus* | -1.231\*\*\* |  | -1.201\*\*\* |  |  |  |
| (financially supported study) | (0.362) |  | (0.395) |  |  |  |
| *robustness\_check* | -0.716\*\* |  |  | -2.915\*\*\* |  |  |
| (robustness checking) | (0.325) |  |  | (0.638) |  |  |
| *psm* |  |  | -2.024\*\* |  |  | -13.36\*\*\* |
| (PSM method) |  |  | (0.814) |  |  | (3.283) |
| \_*cons* | 2.190\*\*\* | 3.714\*\*\* | 4.022\*\*\* | 1.397\*\*\* | 0.584\*\* | 13.19\*\*\* |
| (constant) | (0.734) | (0.679) | (0.681) | (0.299) | (0.229) | (3.580) |
| Observations | 289 | 289 | 357 | 357 | 200 | 200 |
| R-squared | 0.601 | 0.212 | 0.577 | 0.525 | 0.722 | 0.308 |
| Number of studies (clusters) |  | 23 |  | 27 |  | 16 |
| Ramsey RESET  (p-value for H0: linear functional form) | p=0.086 | n.a. | p=0.000 | n.a. | p=0.337 | n.a. |

Notes. Robust standard errors in parentheses. \*\*\*, \*\* and \* denote statistical significance at the 1%, 5% and 10% levels, respectively. n.a. – not applicable

The moderator variables reveal sources of heterogeneity in the effects reported in the literature. Although the post-1999 literature analysed in the present MRA is similar to the pre-2003 literature analysed by the meta-analysis of García-Quevedo (2004) with respect to the heterogeneity of the effects reported, the MRA results contradict García-Quevedo (2004, p.96) which finds that ‘it is not possible to obtain any regularity in the relation between the principal characteristics of the design of applied analyses and the results obtained by them’. To illustrate, we highlight the systematic influence – i.e. across models and subsamples – detected for estimates that do not address the potential endogeneity of subsidies to the selection process (*invSEno\_control\_endogeneity*). In all four cases (all in Input I and Input II) this methodological characteristic is associated with less negative or more positive effects.[[20]](#footnote-20) This is consistent with the emphasis in more recent research on methodologies capable of addressing potential endogeneity and, hence, with our decision to focus on studies published since 2000. Of course, this is now well-known, so that studies not addressing the potential endogeneity of subsidy are unlikely to be published. (In our MRA database, few reported effects arise from methodologies that do not explicitly address this issue; see Table 6.) However, our moderator variables point to other research practices that potentially influence the incidence of omitted variable(s) bias: in particular, difference-in-differences (DiD) estimation (*invSEdid*); propensity score matching (PSM) (*invSEpsm*); and the degree of sample homogeneity (*invSErdperformersonly*). First, we discuss these moderators in the context of other, related influences. Other moderators are discussed where the estimates reveal interesting features of the literature.

Four from five statistically significant estimates of the effect of using panel data (*invSEpanel*) suggest a positive impact on estimated effect size. Taken in conjunction with the weaker evidence of a negative effect from using DiD estimation (*invSEdid*), we suggest that whereas the use of richer data enables larger effects to be identified (possibly due to the ability to capture cumulative effects over time) the use of a panel estimator designed to identify causal treatment effects may tend to reduce the estimated effect size. The positive effects on reported effect size of using more recent data (*invSEmid*\_*point*) are consistent either with increasing effectiveness over time of subsidies (perhaps associated with learning effects by programme managers) or with changing research practices – not otherwise controlled for – being better able to detect unchanging effects. We favour the former explanation, for two reasons: first, because year of publication (*invSEyearofpublication*) is controlled for separately; and, secondly, because we control for a wide variety of research practices. In particular, in three from four significant estimates, across all three subsamples, PSM methods (*invSEpsm*) have a positive effect on reported effects in comparison to other methodological approaches, possibly because PSM controls only for observable characteristics and so may yield estimates potentially affected by omitted variable bias. In contrast, DiD may yield a negative influence, because it takes into account both observable and unobservable firm or plant characteristics, both of which may influence participation in subsidy schemes. Similarly, there is some evidence (two from three significant estimates) that samples including only R&D performers (*invSErdperformersonly*) yield more negative (less positive) estimated effects. This effect may arise because such a sample makes it more likely that the treatment and comparison firms are as similar as possible in all respects other than receipt of subsidy, thereby reducing the effect of unobservable heterogeneity and minimising potential omitted variable bias. Together, these findings on the impact of research practices related to unobservable heterogeneity are consistent with the proposition advanced by Hujer and Radic (2005) and Greene (2009) that evaluation methods taking into account only observable firm characteristics (such as matching methods) yield larger programme effects than those methods controlling further for unobservable influences.

The other moderator variables capturing sample heterogeneity with some degree of consistency across the columns of Table 7 – *invSEeu* and *invSEmanufacturing* – have a control function but are consistent with competing explanations. For example, the positive effect of a manufacturing only sample may conceivably reflect greater financial constraints on private R&D or greater ease of monitoring compared to other sectors, thus rendering crowding out less likely. Among the remaining methodological features, the negative influence on the reported effects associated with using a binary indicator of subsidy (*invSEbinary*) offers the minimal but nonetheless important information that the typical situation confronting researchers – of the amount of subsidy being unknown – does *not* bias estimates in an upward direction.

Finally, two of the *K*-variables reported in Table 7 provide insights into the nature of publication bias in this literature and how it may be reduced in practice.[[21]](#footnote-21) First, the two large, significant and negative estimated coefficients on *financialsupportstatus* suggests that financially-supported studies typically *reduce* publication bias in this literature. Because it is often unclear whether or not an acknowledged funding body has a material interest in the results of the study, this variable most likely aggregates funders with strictly scientific interests and funders with a material stake in the outcome. Consequently, this variable contains a high ratio of noise to signal. Hence, while the hypothesised effect of funders’ influence is positive, the estimate is likely to be statistically insignificant. Conversely, finding a highly significant negative effect at least suggests that this literature is not biased by the perceived interests of funders. Secondly, while robustness checking as a *Z*-variable (*invSErobustness\_check*) is associated with larger reported effect sizes, as a *K*-variable (*robustness\_check*) its negative effect suggests that the practice of reporting results from alternative specifications and estimators counteracts selectivity and thus mitigates publication bias in this literature.

## 6.3 Average genuine empirical effects beyond publication bias

Finally, we estimate the *average* magnitudes of publication selection bias and genuine effects under each of our six multivariate specifications: first, we calculate the linear combination of the constant term (\_cons) and each estimated *K*-effect weighted by its mean value to obtain the representative publication bias; second, we calculate the linear combination of the inverse standard error of the PCC (*invsepcc*) and each *Z*-effect weighted by its mean value to obtain the representative effect size beyond publication bias (Stanley and Doucouliagos, 2012, p.98). Table 8 below reports the results of these calculations.

These findings are consistent with the bivariate MRA estimates reported in section 5. Following the guidelines of Doucouliagos and Stanley (2013), of the six bivariate estimates half suggest “little to modest” publication bias (FAT<1, of which two are statistically insignificant) and half “substantial” publication bias (1≤FAT≤2) (Table 4); while of the six multivariate estimates two suggest “little to modest” publication bias (of which one is not significant), three “substantial” and one “severe” publication bias (FAT>2) (Table 8). Typically, the results reported by this literature are “substantially” biased in a positive direction by publication selection. Once this bias is controlled for, both bivariate and multivariate MRA consistently report positive authentic empirical effects that according to the guidelines of Doucouliagos (2011) (see Footnote 12) can be characterised as either “small” (PET<0.07, 4 bivariate and 4 multivariate estimates) or “moderate” (0.07≤PET≤0.33, 2 bivariate and 2 multivariate).

**Table 8. *Average* publication selection bias and *average* genuine empirical effect calculated for each subsample (derived from multiple MRA; Table 7)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Subsample | Input I | | Input II | | Output | |
| Estimated Model | **WLS MRA**  **(eq. 10)** | **Fixed Effects**  **(eq. 11)** | **WLS MRA**  **(eq. 10)** | **Fixed Effects**  **(eq. 11)** | **WLS MRA**  **(eq. 10)** | **Fixed**  **Effects**  **(eq. 11)** |
| *Average* publication selection bias | 1.17\*\*  [0.03, 2.32]  (t=2.13) | 1.32\*\*\*  [0.47, 2.18]  (t=3.21) | 1.58\*\*\*  [0.86, 2.29]  (t=4.54) | 0.48  [-0.18, 1.14]  (t=1.51) | 0.58\*\*  [0.09, 1.07]  (t=2.54) | 2.09\*\*  [0.20, 3.98]  (t=2.36) |
| *Average* genuine empirical effect | 0.018  [-0.017, 0.054]  (t=1.07) | 0.040 \*\*\*  [0.017, 0.064]  (t=3.58) | 0.082\*\*\*  [0.059, 0.105]  (t=7.25) | 0.103\*\*\*  [0.079, 0.127]  (t=8.66) | 0.066\*\*\*  [0.051, 0.081]  (t=9.24) | 0.061\*\*\*  [0.028, 0.093]  (t=4.01) |

Notes. 95% confidence intervals in square brackets. \*\*\* and \*\* denote statistical significance at the 1% and 5% levels respectively.

The purpose of MRA is not just to reveal publication bias, which is found in most econometric literatures (Doucouliagos and Stanley, 2013), but to identify genuine empirical effects “beyond publication bias”. Once controlled for, publication bias has no further implications for interpreting the effects of subsidies. As to why publication bias might exist, although our *K*-moderator variables reveal some influences on publication bias, they cannot account for influences that affect all researchers contributing to the literature.[[22]](#footnote-22) Here, we have to move beyond our data and our remarks are accordingly tentative. If a source of publication bias is conformity to some mainstream theory then a corollary of competing theories should be reduced publication bias, which is a conjecture with some empirical support (Doucouliagos and Stanley, 2013). Yet, in the case of the literature under consideration, the theoretical reasons for contrasting findings outlined in Section 2 can scarcely be described as competing: on the one hand, market failures well-known and hallowed by constant reiteration suggest a role for R&D subsidy (i.e. the “mainstream perspective” reviewed in Section 2 above); on the other hand are relatively unknown suggestions as to why subsidy might in practice fail (i.e. the “public choice” and “asymmetric information” perspectives that we advance in Section 2). Reflecting this asymmetry, we conjecture that any bias in the research process will be in a positive direction, towards additionality or at least against crowding out. Of course, the competing possibilities of additionality and crowding out are also long established and may temper such bias, which may explain why publication bias is “substantial” but not “severe”.

## 6.4 Out-of-sample prediction

If our MRA models are a valid representation of the literature under consideration then they should be able to predict the R&D effects reported in studies of the same type – i.e. part of the same population – but that are not included in the estimation sample, so long as these out-of-sample reported estimates are not the result of selection bias. Accordingly, successful out-of-sample prediction provides a joint check on the validity of our MRA models, the consistency of the out-of-sample studies with the studies in the MRA database, and the absence of extreme selection bias in the out-of-sample estimates.[[23]](#footnote-23) We use two studies published after our cut-off point.[[24]](#footnote-24) First, we code the estimates reported in these studies together with their sample and research characteristics in the same way as we code the 52 studies in our MRA database; secondly, we transform the reported effects into PCCs; then we substitute the values of the corresponding moderator variables for each of these two studies into the multiple MRA models reported in Table 7 to derive the predicted PCC values. The results of this exercise are reported in Table 9 below.

**Table 9. MRA “out-of-sample” predictions**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Studies** | **Subsample** | **Out-of-sample reported effect**  **(transformed into PCCs)** | **Average genuine effect in the literature**  **(MRA estimates from Table 8)** | **Out-of-sample prediction**  **[95% Confidence Interval]** |
| **Hud**  **& Hussinger (2014)** | Input I | 0.087\*\*\* | **WLS** | **WLS** |
| 0.018  [-0.017, 0.054] | 0.053\*\*  [0.007, 0.099] |
| **FE** | **FE** |
| 0.040\*\*\*  [0.017, 0.064] | 0.088\*\*  [0.019, 0.158] |
| Input II | 0.087\*\*\* | **WLS** | **WLS** |
| 0.082\*\*\*  [0.059, 0.105] | 0.022\*\*  [0.005, 0.038] |
| **FE** | **FE** |
| 0.103\*\*\*  [0.079, 0.127] | 0.066\*\*\*  [0.036, 0.097] |
| **Hottenrott**  **& Lopes-Bento (2014)** | Input II | 0.123\*\*\* | **WLS** | **WLS** |
| 0.082\*\*\*  [0.059, 0.105] | 0.087\*\*\*  [0.029, 0.146] |
| **FE** | **FE** |
| 0.103\*\*\*  [0.079, 0.127] | 0.055\*\*\*  [0.028, 0.082] |

Notes. 95% confidence intervals in square brackets. \*\*\* and \*\* denote statistical significance at the 1% and 5% levels respectively.

Hud and Hussinger (2014) report effects on both private R&D expenditure (Input I) and total R&D expenditure (Input II), in both cases yielding PCCs of 0.087; and Hottenrott and Lopes-Bento (2014) report an estimate only for total R&D expenditure (Input II), yielding a PCC of 0.123.[[25]](#footnote-25) The next two columns display, respectively, the “average genuine effects” estimated by multiple MRA from our sample studies (reproduced from Table 8) and the effects predicted by our MRA models for these two studies. In each case the predicted effects are statistically significant and positive, so qualitatively the model performs uniformly well. Quantitatively, each of the three reported effects falls within the 95 per cent confidence intervals around the effects predicted by at least one of our MRA models: both WLS and FE predictions for the Hud and Hussinger (2014) Input I effect; FE for the Hud and Hussinger (2014) Input II effect; and WLS for the Hottenrott and Lopes-Bento (2014) Input II effect. However, the two reported effects that do not lie within the confidence intervals of the corresponding prediction both lie above and close to the upper confidence interval. This is consistent with the MRA predictions being net of positive publication bias and so reduced in size. Accordingly, four from six quantitative comparisons suggest no significant differences at the five per cent level between the effects reported in these two studies and MRA model predictions corrected for publication bias, while the remaining two yield differences so small as to be of no economic significance.

We assume that these out-of-sample studies are from the same population as our MRA sample, given that no new coding categories were required to capture their respective research practices or context. In this case, failure to predict out-of-sample could result either from the out-of-sample studies reporting estimates subject to selection bias or/and lack of validity of the MRA model(s) used to make the predictions. Conversely, the results of this particular exercise suggest both that the out-of-sample studies are not greatly affected by selection bias and that our MRA models are valid representations of the literature under consideration.

# Concluding remarks

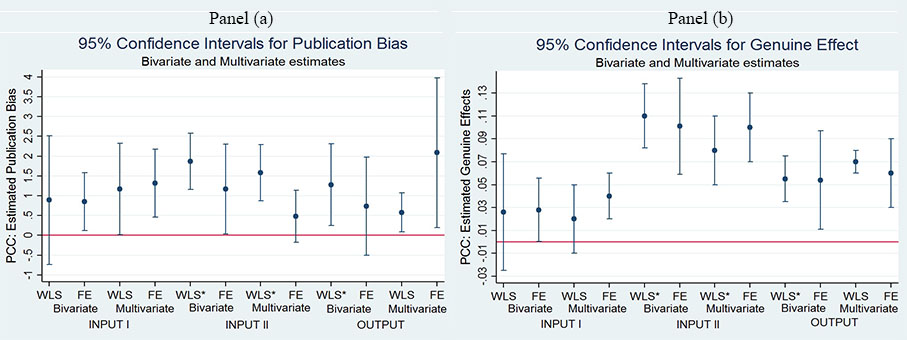
For the best part of half a century the developed economies have used subsidies to promote R&D. Yet, neither theory nor the extensive empirical literature yield unambiguous conclusions as to the size or even the direction of the effect of public R&D funding on private R&D. Accordingly, we conduct a meta-regression analysis of the literature since 2000 – comprising 52 primary studies – to identify the genuine representative effect established by this literature after controlling for publication bias and sources of heterogeneity.

We analyse a broad range of studies evaluating the subsidy effects of two different measures of subsidy inputs (binary and level) on diverse R&D inputs and R&D outputs. This heterogeneity is reflected in our division of the literature into three subsamples. Moreover, there is no standard effect size in the literature. Accordingly, we follow best practice in MRA by transforming the heterogeneous effects reported in the literature into partial correlation coefficients (PCCs) between R&D subsidies and R&D outcomes.

This MRA applies methods to identify publication bias and to control for it in estimating the genuine empirical effect beyond publication bias. This is a contribution that complements not only narrative review, which is less amenable to reproduction and verification, but also simple meta-analysis, which neither controls for publication bias nor estimates a genuine representative effect beyond publication bias. A novel feature of this MRA is “out-of-sample” prediction.

Figure 5 provides a graphical summary of our main results: for each of the four models (two bivariate and two multivariate), estimated for each of our three subsamples, we display both the estimated publication bias (Panel a) and the estimated genuine empirical effect beyond – or corrected for – publication bias (Panel b) identified by this literature. Estimates that are not statistically significant at the five per cent level are those with confidence intervals that include zero; and those failing the standard diagnostic test for linearity are denoted with an asterisk.

Figure 5. Summary of MRA estimates of publication bias (Panel a) and genuine empirical effects (Panel b)



Notes. \* Denotes diagnostic failure at the 1% significance level with respect to the hypothesised linear functional form.

Estimates of the publication bias in this literature are consistent both within and between our subsamples: within each group, the statistically significant results all have confidence intervals that overlap (so that variations between estimates are not significant); and between groups, with one exception the significant estimates overlap. By standards commonly applied in MRA, this positive publication bias is “substantial”, which may reflect the asymmetric weight of theory in this literature: on the one hand, well-known market failures supporting R&D subsidy; on the other, relatively unknown suggestions as to why subsidy might in practice fail. In any case, estimation by meta-regression of the genuine effect identified by this literature controls for, hence, nets out, publication bias.

Estimates of the genuine empirical effect mainly indicate a “small” positive effect. However, the meaning of these results is somewhat different according to the subsample considered.

* Statistically significant positive estimates within the Input I subsample indicate additionality, while insignificant estimates indicate no effect. However, both exclude crowding out of any degree.
* Statistically significant positive estimates within both the Input II and the Output subsamples indicate the exclusion of full and over-full crowding out.

The PCC is a standard measure but not one of economic effect. To provide a direct measure of subsidy effectiveness, we analyse a further subsample of constant elasticities (the percentage response of firms’ R&D spending to a percentage change in subsidy), which do not require transformation. Although the sample is small, the models are diagnostically satisfactory and yield estimated elasticities of less than 0.01 which, although statistically significant, are economically negligible. Indicatively, a doubling of subsidy would yield an increase in private R&D of less than one per cent.

In sum, *MRA findings reject crowding out of private investment by public subsidy but reveal no evidence of substantial additionality*. This conclusion contributes to policy debate by identifying a representative subsidy effect from the large and complex evaluation literature. Our findings suggest that the use of subsidy as part of science and technology policy does contribute to addressing market failures by increasing both R&D inputs and R&D outputs in subsidised firms in comparison to the no subsidy counterfactual. We find that subsidies are generally not wasted; this literature rejects complete crowding out. This could be especially important when R&D subsidies are implemented as part of a wider counter-cyclical policy in order to sustain R&D investment during economic downturns, as recent research indicates (Hud and Hussinger, 2014).

While lack of evidence for substantial additionality may be disappointing for policy makers, we suggest that this may be typical of public policies: individual policies can work in the direction intended but yield quantitatively smaller effects than hoped for.[[26]](#footnote-26) However, we find that subsidy effectiveness is increasing over time, which may reflect institutional learning. In any case, the overall impact of subsidies is underestimated by this evaluation literature to the extent that the potential substitutability of input and output additionalities with behavioural additionality (Georghiou, 2002; Clarysse et al., 2009) is not accounted for; subsidised firms might benefit from uncaptured behavioural effects. If so, the lack of substantial additionality identified in this MRA may not fully capture the effects of R&D subsidy programmes on subsidised firms. This possibility is consistent with the positive influence on estimated effect size of using panel data to estimate subsidy effects, because longitudinal data is necessary to capture behavioural effects manifested over time. Accordingly, we endorse Falk (2006) who, in addition to investigating both behavioural and input additionality, emphasises the need for better measures and data to evaluate behavioural additionality.

Our results also have implications for research practice and the interpretation of findings in this literature. Multiple MRA contradicts the finding of the earlier meta-analysis (García-Quevedo, 2004) by revealing relationships between research practices and heterogeneity in the effects reported by the primary literature. First, we find that estimates failing to address the potential endogeneity of subsidy tend to report larger effects. Of course, it is now accepted practice to address the possibility that firms’ propensity to receive subsidy is related to their observable or unobservable characteristics, in order to mitigate the influence of omitted variable(s) bias on estimated subsidy effects. However, our findings suggest that other research practices may also influence the incidence of omitted variable(s) bias: in particular, difference-in-differences (DiD) estimation tends to reduce while propensity score matching (PSM) tends to increase the size of estimated effects; and sample homogeneity tends to reduce the size of reported effects. We argue that the common thread between these findings is the influence of unobservable firm or plant characteristics which, if positively related to subsidy receipt, may positively bias estimated subsidy effects. Each of these research practices has implications for the effect of unobservable characteristics: DiD controls for unobservable influences, which eliminates a source of positive bias; PSM does not control for unobservable influences, which permits a potential source of positive bias; and homogeneous samples reduce potential differences in unobservable characteristics between treatment and comparison groups, thereby attenuating potential bias. These findings from multiple MRA thus provide evidence consistent with the conjecture of Hujer and Radic (2005) and Greene (2009) that evaluation methods controlling for unobservable influences find smaller programme effects than do methods controlling only for observable influences. Two other findings are also noteworthy from the perspective of research practice in this literature: using a binary indicator of subsidy – the typical practice in this literature – does *not* bias estimates in an upward direction; and robustness checking – the practice of reporting results from alternative specifications and estimators – mitigates publication bias.

# Acknowledgements

The authors are grateful for many helpful suggestions from participants at the 2013 Meta-Analysis in Economic Research Colloquium at the University of Greenwich. Dragana Radicic and Mehmet Ugur were especially generous in providing timely feedback.

**References**

**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, 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.

**Bloom, N., Schankerman, M., Van Reenen, J. (2013)** Identifying technology spillovers and product market rivalry, Econometrica 81(4), 1347-1393.

**Busom, I. (2000)** An empirical evaluation of R&D subsidies, Economics of Innovation and New Technology 9(2), 111-148.

**Butler, E. (2012)** Public Choice - A Primer, London: Institute of Economic Affairs and Profile Books. Accessed 22-05-2015: <http://www.iea.org.uk/sites/default/files/publications/files/IEA%20Public%20Choice%20web%20complete%2029.1.12.pdf>

**Cerulli, G. (2010)** Modelling and measuring the effect of public subsidies on business R&D: a critical review of the econometric literature, The Economic Record 86, 421-449.

**Clarysse, B., Wright, M., Mustar, P. (2009)** Behavioural additionality of R&D subsidies: A learning perspective, Research Policy 38 (10), 1517–1533.

**Clausen, T. H. (2009)** Do subsidies have positive impacts on R&D and innovation activities at the firm level?, Structural Change and Economic Dynamics 20(4), 239-253.

**Cooper, H.M, Hedges, L.V. (eds.) (1994)** Handbook of Research Synthesis, New York: Russell Sage.

**Czarnitzki, D. (2006)** Research and Development in Small and Medium-Sized Enterprises: The Role of Financial Constraints and Public Funding, Scottish Journal of Political Economy 53(3), 335-357.

**Czarnitzki, D., Hottenrott, H., Thorwarth, S. (2011)** Industrial research versus development investment: the implications of financial constraints, Cambridge Journal of Economics 35(3), 527-544.

**David, P. A., Hall, B. H., Toole, A. A. (2000)** Is public R&D a complement or substitute for private R&D? A review of the econometric evidence, Research Policy 29, 497–529.

**Doucouliagos, H. (2011)** How Large is Large? Preliminary and relative guidelines for interpreting partial correlations in economics, School Working Paper of Economics Series, Deakin University.

**Doucouliagos, H., Stanley, T.D. (2009)** Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis, British Journal of Industrial Relations 47(2), 406-428.

**Doucouliagos, H., Stanley, T.D. (2013)** Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity, Journal of Economic Surveys 27(2), 316-339.

**Falk R. (2006)** Behavioural additionality of Austria’s industrial research promotion fund (FFF) in ‘Government R&D funding and company behaviour: Measuring behavioural additionality’, OECD, 59-74.

**García-Quevedo, J. (2004)** Do Public Subsidies Complement Business R&D? A Meta-Analysis of the Econometric Evidence, Kyklos 57(1), 87-102.

**Georghiou, L. (2002)** Additionality and impact of R&D subsidies, IWT Studies 40, 57-64.

**Greene. F. J. ( 2009)** Assessing the impact of policy interventions: the influence of evaluation methodology, Environment and Planning C: Government and Policy 27(2), 216-229.

**Hall, B. (2002a)** The Assessment: Technology Policy, Oxford Review of Economic Policy 18, 1-9.

**Hall, B. (2002b)** The Financing of Research and Development, Oxford Review of Economic Policy 18, 35-51.

**Hottenrott, H., Lopes Bento, C. (2014)** (International) R&D Collaboration and SMEs: The Effectiveness of Targeted Public R&D Support Schemes, Research Policy 43(6), 1055-1066.

**Hud, M., Hussinger, K. (2014)** The Impact of R&D Subsidies During the Crisis, ZEW Discussion Paper No. 14-024, Mannheim.

**Hujer, R., Radic, D. (2005)** Evaluating the Impacts on Innovation Activities in Germany, Scottish Journal of Political Economy 52(4), 565-586.

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

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

**Moreno, S.G., Sutton, A.J., Ades, A., Stanley, T.D., Abrams, K.R., Peters, J.L., Cooper, N.J. (2009)** Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study, BMC Medical Research Methodology 9(2).

Open access: <http://www.biomedcentral.com/1471-2288/9/2>

**Nelson, R. R. (1959)** The Simple Economics of Basic Scientific Research, Journal of Political Economy 67, 297-306.

**OECD** **(2013)** OECD Science, Technology and Industry Scoreboard 2013. Accessed: 22-05-2015: [www.oecd-ilibrary.org/docserver/download/9213051e.pdf](file:///G:\2_M%20R%20A%20%20%20P%20A%20P%20E%20R\MRA_resubmission\Article_resubmission\www.oecd-ilibrary.org\docserver\download\9213051e.pdf)

**Ozcelik, E., Taymaz, E. (2008)** R&D support programs in developing countries: The Turkish experience, Research Policy 37, 258–275.

**Pugh, G., Haile, M. and Ćorić, B (2011a)** An Introduction to Meta-Regression Analysis (MRA): Using the Example of the Trade Effects of Exchange Rate Variability, in Guido Erreygers and Mieke Vermeire (eds), Macroeconomics and Beyond: Essays in honour of Wim Meeusen (Antwerp-Apeldoorn: Garant).

**Pugh, G., Mangan, J., Gray, J. (2011b)** Do increased resources increase educational attainment during a period of rising expenditure? Evidence from English secondary schools using a dynamic panel analysis, British Educational Research Journal, Volume 37(1), 163-189.

**Pugh, G., Mangan, J., Blackburne, V., Radicic, D. (2014)** School Expenditure and School Performance: Evidence from New South Wales schools using a dynamic panel analysis, British Educational Research Journal, Vol. 41(2), 244-264.

**Radicic, D., Pugh, G., Hollanders, H., Wintjes, R. (2014)** The impact of innovation support programmes on SME innovation in traditional manufacturing industries: an evaluation for seven EU regions, Maastricht Economic and Social Research institute on Innovation and Technology (UNU‐MERIT), Working Paper #2014-033. Accessed 22-05-2015: <http://pub.maastrichtuniversity.nl/d58cc147-c79b-4529-ac1f-e87d8bbb88d1>

**Radicic, D., Pugh, G., Hollanders, H., Wintjes, R. (forthcoming)** The impact of innovation support programmes on SME innovation in traditional manufacturing industries: an evaluation for seven EU regions, Environment and Planning C: Government and Policy.

**Schumpeter, J. A. (1942)** Capitalism, Socialism and Democracy, London: Routledge. pp. 82–83.

**Solow, R. M. (1956)** A Contribution to the Theory of Economic Growth, Quarterly Journal of Economics 70(1), 65 – 94.

**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, 103-27.

**Stanley, T.D., Doucouliagos, C. (H.) (2007)** Identifying and correcting publication selection bias in the efficiency-wage literature: Heckman meta-regression, School Working Paper, Economics Series 2007-11, Deakin University.

**Stanley, T.D., Doucouliagos, C. (H.) (2010)** Picture this: A simple graph that reveals much ado about research, Journal of Economic Surveys 24, 170-91.

**Stanley, T.D., Jarrell, S. B., Doucouliagos, H. (2010)** Could It Be Better to Discard 90% of the Data? A Statistical Paradox, The American Statistician 64, 70-77.

**Stanley, T.D., Doucouliagos, C. (H.) (2011)** Meta-regression approximations to reduce publication selection bias, School Working Paper, Economics Series 2011-4, Deakin University.

**Stanley, T. D., Doucouliagos, H. (2012)** Meta-regression analysis in economics and business, Routledge Advances in Research Methods (Book 5), 1st edition, July 2012, Routledge.

**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.

**Usher, D. (1964)** The Welfare Economics of Invention, Econometrica, 31, 279-287.

**Zúñiga-Vicente, J. Á., Alonso-Borrego, C., Forcadell, F. J., Galán, J. I. (2014)** Assessing the effect of public subsidies on firm R&D investment: a survey, Journal of Economic Surveys 28(1), 36-67.

**Appendix A. Compendium Table of coded studies**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Studies** | **Observations**  (range of observations in the study corresponding to each regression/match) | **Type of Data** | | | **Developed Economy** | **Sample Period**  (of coded regressions/matches) | **R&D performers only**  (percentages refer to reported regressions/matches) | **Estimation Method** | **Type of R&D Outcome** | **Measure of Treatment** | **Main results**  (percentages refer to reported regressions/matches) |
| **Cross** | **Panel** | **Pooled** |
| **Aerts & Czarnitzki**  **(2004)** | 350 -676 |  |  | X | Yes | 1998 - 2000 | 50% Yes  50% No | PSM and OLS (no control of endogeneity) | Both Input and Output | Binary | * 85% rejection of full crowding out * 15% full crowding out |
| **Aerts & Schmidt**  **(2008)** | 314 - 968 |  |  | X | Yes | 1998 - 2004 | 43% Yes  57% No | PSM, OLS and Conditional DiD for Repeated Cross Sections | Input Only | Binary | * 81% rejection of full crowding out * 9% full crowding out |
| **Aerts & Thorwarth (2008)** | 521 |  |  | X | Yes | 2002 - 2006 | Yes | Selection model & IV estimation | Input Only | Both binary and continuous | * 67% additionality * 33% no effect |
| **Alecke et al.**  **(2012)** | 225-1237 | X |  |  | Yes | 2003 | No | PSM and OLS (no control of endogeneity) | Both Input and Output | Binary | * 40% rejection of full crowding out * 50% additionality * 10% full crowding out |
| **Ali-Yrkkö**  **(2005)** | 1610 - 1640 |  |  | X | Yes | 1996 - 2002 | No | IV estimation and OLS (no control of endogeneity) | Input Only | Binary | * 87% additionality * 13% no effect |
| **Almus & Czarnitzki (2003)** | 1244 |  |  | X | No | 1995 - 1997 | No | PSM | Input Only | Binary | * 100% additionality |
| **Aschhoff**  **(2009)** | 282 - 3942 |  |  | X | Yes | 1994 - 2005 | Yes | PSM and OLS | Both Input and Output | Both binary and continuous | * 41% rejection of full crowding out * 34% additionality * 19% no effect * 6% full crowding out |
| **Aschhoff & Fier (2005)** | 606 |  |  | X | Yes | 1996 - 2003 | No | PSM | Both Input and Output | Binary | * 100% rejection of full crowding out |
| **Bérubé & Mohnen (2009)** | 1066 - 1168 |  |  | X | Yes | 2002 -2004 | No | PSM | Output | Binary | * 81% rejection of full crowding out * 19% full crowding out |
| **Bloch & Graversen (2012)** | 450 |  | X |  | Yes | 1997 - 2005 | No | OLS and Selection Model (IV estimation) | Input | Continuous | * 100% additionality |
| **Busom**  **(2000)** | 143 | X |  |  | Yes | 1988 | Yes | Selection model | Input | Binary | * 100% rejection of full crowding out |
| **Carboni**  **(2011)** | 457 - 1235 |  |  | X | Yes | 2001 - 2003 | No | PSM and OLS (no control of endogeneity) | Input | Binary | * 71% additionality * 29% no effect |
| **Cerulli & Potì**  **(2012)** | 64 - 5097 |  | X | X | Yes | 1998 - 2003 | No | PSM, Selection model, DiD and OLS (no control of endogeneity) | Input | Binary | * 84% rejection of full crowding out * 16% full crowding out |
| **Clausen**  **(2009)** | 1019 |  |  | X | Yes | 1998 - 2001 | Yes | IV estimation | Input | Continuous | * 12% rejection of full crowding out * 25% additionality * 38% no effect * 13% partial crowding out * 6% full crowding out * 6% over-full crowding out |
| **Czarnitzki et al. (2007)** | 202 - 992 |  |  | X | Yes | 1994 - 1999 | No | PSM | Both Input and Output | Binary | * 58% rejection of full crowding out * 42% full crowding out |
| **Czarnitzki & Fier (2001)** | 2451 |  |  | X | Yes | 1994 - 1998 | No | OLS (subsidy variable is lagged, so considered not to be endogenous) | Input | Continuous | * 100% additionality |
| **Czarnitzki & Fier (2002)** | 420 |  |  | X | Yes | 1996 - 1997 | No | PSM | Input | Binary | * 100% rejection of full crowding out |
| **Czarnitzki et al. (2011)** | 839 - 1085 |  | X |  | Yes | 1999 - 2007 | No | Control function (panel analysis) | Input | Binary | * 25% rejection of full crowding out * 75% full crowding out |
| **Czarnitzki**  **& Hussinger**  **(2004)** | 468 - 3764 |  |  | X | Yes | 1992 - 2000 | Yes | PSM and OLS (after having calculated the subsidy induced R&D) | Both Input and Output | Binary | * 67% rejection of full crowding out * 33% additionality |
| **Czarnitzki & Licht (2006)** | 491 - 1462 |  |  | X | 50% Yes  50% No | 1994 - 1998 | 42% Yes  58% No | PSM and OLS (after having calculated the subsidy induced R&D) | Both Input and Output | Binary | * 96% rejection of full crowding out * 4% full crowding out |
| **Czarnitzki**  **& Lopes-Bento**  **(2011)** | 235 - 3016 |  |  | X | Yes | 1992 - 2005 | No | PSM | Both Input and Output | Binary | * 56% rejection of full crowding out * 44% full crowding out |
| **Czarnitzki**  **& Lopes-Bento**  **(2012)** | 126 - 6006 |  |  | X | 67% Yes  33% No | 2002 - 2004 | No | PSM and OLS (no control of endogeneity) | Input | Binary | * 100% rejection of full crowding out |
| **Czarnitzki**  **& Lopes-Bento**  **(2013)** | 224 - 4761 |  |  | X | Yes | 2002 - 2008 | No | PSM and IV estimation | Input | Binary | * 100% rejection of full crowding out |
| **Czarnitzki & Toole (2007)** | 925 |  |  | X | Yes | 1998 - 1999 | No | OLS (no control of endogeneity)\* | Input | Binary | * 100% rejection of full crowding out |
| **Duguet**  **(2004)** | 512 - 958 |  |  | X | Yes | 1985 - 1997 | Yes | PSM | Input | Binary | * 92% no effect * 8% crowding out of any degree |
| **Dumont**  **(2013)** | 11639 |  |  | X | Yes | 2001 - 2009 | No | Selection model | Input | Continuous | * 100% additionality |
| **Ebersberger**  **(2005)** | 104 - 1575 | X |  | X | Yes | 1994 - 2000 | No | PSM, Selection model, OLS (no control of endogeneity) and quasi DiD | Both Input and Output | Binary | * 78% rejection of full crowding out * 22% full crowding out |
| **Ebersberger**  **& Lehtoranta**  **(2008)** | 606 |  |  | X | Yes | 2002 - 2004 | No | PSM | Both Input and Output | Binary | * 57% rejection of full crowding out * 43% full crowding out |
| **Einio**  **(2009)** | 752 - 1656 |  | X |  | Yes | 2000 - 2005 | No | IV estimation and OLS (subsidy variable is lagged, so considered not to be endogenous) | Input | Binary | * 89% rejection of full crowding out * 11% full crowding out |
| **Falk**  **(2006)** | 638 - 3031 |  | X |  | Yes | 1995 - 2002 | No | Control Function (Panel analysis) | Input | Continuous | * 87% rejection of full crowding out * 13% full crowding out |
| **González et al.**  **(2005)** | 5076 |  |  | X | Yes | 1990 - 1999 | No | Selection model | Input | Continuous | * 100% additionality |
| **González & Pazó (2008)** | 410 - 1306 |  |  | X | Yes | 1990 -1999 | 45% Yes  55% No | PSM | Input | Binary | * 45% rejection of full crowding out * 23% additionality * 32% no effect |
| **Görg & Strobl**  **(2007)** | 144 - 4192 |  | X | X | Yes | 1999 - 2002 | No | DiD & Matching, DiD and OLS (no control of endogeneity) | Input | Continuous | * 21% additionality * 65% no effect * 7% partial crowding out * 7% over-full crowding out |
| **Halpern**  **& Muraközy**  **(2010)** | 6514 |  |  | X | No | 2002 - 2006 | Yes | Crepon-Duguet-Mairesse (CDM) model | Both Input and Output | Binary | * 100% rejection of full crowding out |
| **Heijs & Herrera (2004)** | 88 - 366 |  |  | X | Yes | 1998 - 2000 | No | PSM | Input | Continuous | * 100% partial crowding out |
| **Herrera**  **& Bravo Ibarra**  **(2010)** | 135 - 334 | X |  |  | Yes | 1999 - 2001 | No | PSM | Both Input and Output | Binary | * 17% rejection of full crowding out * 67% additionality * 8% no effect * 8% full crowding out |
| **Herrera et al.**  **(2007)** | 150 - 416 |  |  | X | Yes | 1999 - 2000 | No | PSM | Both Input and Output | Binary | * 92% rejection of full crowding out * 8% full crowding out |
| **Hewitt-Dundas**  **& Roper**  **(2010)** | 656 - 1007 |  |  | X | Yes | 1994 - 2003 | No | IV estimation | Output | Binary | * 50% rejection of full crowding out * 50% full crowding out |
| **Hujer & Radić**  **(2005)** | 134 - 2013 |  | X | X | 79% Yes  21% No | 1997 - 2000 | No | PSM, Selection model, OLS (no control of endogeneity) and Conditional DiD | Output | Binary and Continuous (R&D induced by public funding) | * 46% rejection of full crowding out * 44% full crowding out * 10% over-full crowding out |
| **Hussinger**  **(2008)** | 2129 - 3744 |  |  | X | Yes | 1992 - 2000 | No | Selection model and OLS (no control of endogeneity) | Both Input and Output | Binary | * 56% rejection of full crowding out * 44% rejection of full crowding out |
| **Kaiser**  **(2004)** | 68 - 1134 |  |  | X | Yes | 1999 - 2001 | No | PSM and Selection model | Input | Binary | * 73% no effect * 27% crowding out of any degree |
| **Klette**  **& Moen**  **(2012)** | 181 - 697 |  | X |  | Yes | 1985 - 1993 | Yes | Control Function (Panel analysis) | Input | Continuous | * 67% additionality * 33% no effect |
| **Koga**  **(2005)** | 196 - 642 |  | X |  | Yes | 1995 - 1998 | No | Control Function (Panel analysis) | Input | Binary | * 67% additionality * 33% no effect |
| **Koski**  **(2008)** | 1122 |  |  | X | Yes | 1999 - 2003 | No | Selection model | Output | Continuous | * 50% rejection of full crowding out * 50% full crowding out |
| **Lach**  **(2002)** | 39 - 325 |  | X |  | Yes | 1990 -1995 | Yes | DiD | Input | Binary and Continuous | * 6% additionality * 59% no effect * 35% over-full crowding out |
| **Lööf**  **& Heshmati**  **(2005)** | 156 - 462 |  |  | X | Yes | 1998 - 2000 | Yes | PSM | Input | Binary | * 100% rejection of full crowding out |
| **Meeusen**  **& Janssens**  **(2001)** | 36 - 689 |  | X |  | Yes | 1992 - 1997 | Yes | Control Function (Panel analysis) | Input | Continuous | * 90% additionality * 10% no effect |
| **Ozcelik**  **& Taymaz**  **(2008)** | 40 - 98366 |  | X |  | No | 1993 - 2001 | 62% Yes  38% No | IV estimation, DiD, DiD & Matching, Control Function (Panel analysis) | Input | Binary | * 62% rejection of full crowding out * 28% additionality * 10% no effect |
| **Reinkowski et al. (2010)** | 226 - 952 | X |  |  | No | 2003 | 4% Yes  96% No | PSM | Both Input and Output | Binary | * 88% rejection of full crowding out * 12% full crowding out |
| **Streicher et al.**  **(2004)** | 65 - 2194 |  | X |  | Yes | 1997 - 2002 | 63% Yes  37% No | Control Function (Panel analysis) | Input | Continuous | * 100% additionality |
| **Suetens**  **(2002)** | 1032 |  |  | X | Yes | 1992 - 1999 | Yes | IV estimation and OLS (no control of endogeneity) | Input | Continuous | * 75% additionality * 25% no effect |
| **Wallsten**  **(2000)** | 81 - 481 |  |  | X | Yes | 1991 - 1993 | No | IV estimation and OLS (no control of endogeneity) | Input | Continuous | * 33% rejection of full crowding out * 33% partial crowding out * 33% full crowding out |

\* According to the authors, when endogeneity is accounted for the reported results ‘continue to hold’ (Czarnitzki and Toole, 2007, p.179).

Key: PSM stands for Propensity Score Matching; OLS for Ordinary Least Squares; DiD for Difference-in-Differences; IV for Instrumental Variable.

**Studies in the MRA database**

* **Aerts, K., Czarnitzki, D. (2004)** Using Innovation Survey Data to Evaluate R&D Policy: The Case of Belgium, ZEW Discussion Papers 04-55, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.
* **Aerts, K., Schmidt T. (2008)** Two for the price of one? Additionality effects of R&D subsidies: A comparison between Flanders and Germany, Research Policy 37(5), 806-822.
* **Aerts, K., Thorwarth, S. (2008)** Additionality effects of public R&D funding: ‘R’ versus ‘D’, Open Access publications from Katholieke Universiteit Leuven, urn:hdl:123456789/216618, Katholieke Universiteit Leuven.
* **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.
* **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, Journal of Business and Economic Statistics 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.
* **Aschhoff, B., Fier, A. (2005)** Powerful or Powerless? The impact of public R&D grants on SMEs in Germany, Centre for European Economic Research (ZEW), Mannheim.
* **Bérubé, C., Mohnen, P. (2009)** Are firms that receive R&D subsidies more innovative?, Canadian Journal of Economics 42(1), 206-225.
* **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.
* **Busom, I. (2000)** An empirical evaluation of R&D subsidies, Economics of Innovation and New Technology 9(2), 111-148.
* **Carboni, O. (2011)** R&D subsidies and private R&D expenditures: evidence from Italian manufacturing data, International Review of Applied Economics 25(4), 419-439.
* **Cerulli, G., Potì, B. (2012)** Evaluating the robustness of the effect of public subsidies on firms' R&D: an application to Italy, Journal of Applied Economics 15(2), 287-320.
* **Clausen, T. H. (2009)** Do subsidies have positive impacts on R&D and innovation activities at the firm level?, Structural Change and Economic Dynamics 20(4), 239-253.
* **Czarnitzki, D., Ebersberger, B., Fier, A. (2007)** The relationship between R&D collaboration, subsidies and R&D performance: empirical evidence from Finland and Germany, Journal of Applied Econometrics 22(7), 1347-1366.
* **Czarnitzki, D., Fier, A. (2001)** Do R&D subsidies matter? Evidence from the German service sector, ZEW Discussion Paper No.019, Mannheim.
* **Czarnitzki, D., Fier, A. (2002)** Do innovation subsidies crowd out private investment? Evidence from the German service sector, Applied Economics Quarterly 48(1), 1-25.
* **Czarnitzki, D., Hottenrott, H., Thorwarth, S. (2011)** Industrial research versus development investment: the implications of financial constraints, Cambridge Journal of Economics 35(3), 527-544.
* **Czarnitzki, D., Hussinger, K. (2004)** The link between R&D subsidies, R&D spending and technological performance, ZEW Discussion Paper No. 056, Mannheim.
* **Czarnitzki, D., Licht, G. (2006)** Additionality of public R&D grants in a transition economy. The case of Eastern Germany, Economics of Transition 14(1), 101-131.
* **Czarnitzki, D., Lopes-Bento, C. (2011)** Innovation Subsidies: Does the Funding Source Matter for Innovation Intensity and Performance? Empirical Evidence from Germany, ZEW Discussion Paper No. 11-053, Mannheim.
* **Czarnitzki, D., Lopes-Bento, C. (2012)** Evaluation of public R&D policies: A cross-country comparison, World Review of Science, Technology and Sustainable Development 9(2/3/4), 254 - 282.
* **Czarnitzki, D., Lopes-Bento, C. (2013)** Value for money? New microeconometric evidence on public R&D grants in Flanders, Research Policy 42(1), 76-89.
* **Czarnitzki, D., Toole, A. A (2007)** Business R&D and the interplay of R&D subsidies and product market uncertainty, Review of Industrial Organization 31(3), 169-181.
* **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, Michel (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.
* **Ebersberger, Bernd (2005)** The impact of Public R&D Funding, VTT Publications 588, Espoo 2005.
* **Ebersberger, B., Lehtoranta, O. (2008)** Effects of Public R&D Funding, VTT Working Papers 100, Espoo.
* **Einio, E. (2009)** The effect of government subsidies on private R&D: evidence from geographic variation in support program funding, Discussion paper No. 263, ISSN 1795-0562, University of Helsinki and HECER.
* **Falk R. (2006)** Behavioural additionality of Austria’s industrial research promotion fund (FFF) in *Government R&D funding and company behaviour: Measuring behavioural additionality*, OECD, 59-74.
* **González, X., Jaumandreu, J., Pazó, C. (2005)** Barriers to innovation and subsidy effectiveness, RAND Journal of Economics 36(4), 930-950.
* **González, X., Pazó, C. (2008)** Do public subsidies stimulate private R&D spending?, Research 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.
* **Halpern, L., Muraközy, B. (2010)** R&D subsidies and firm performance in Hungary, MICRO-DYN Working Paper no. 38/10.
* **Heijs, J., Herrera, L. (2004)** Distribution of R&D subsidies and its effect on the final outcome of innovation policy, Working paper nº 46, IAIF.
* **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.
* **Herrera, L., Heijs, J., Nieto, M. (2007)** Distribution and effect of R&D subsidies: a comparative analysis according to firm size, Working paper #1463 at DRUID Summer Conference 2007, Denmark.
* **Hewitt-Dundas, N., Roper, S. (2010)** Output additionality of public support for innovation: evidence for Irish manufacturing plants, European Planning Studies 18(1), 107-122.
* **Hujer, R., Radić, D. (2005)** Evaluating the Impacts of Subsidies on Innovation Activities in Germany, ZEW Discussion Papers, No. 05-43.
* **Hussinger, K. (2008)** R&D and subsidies at the firm level: an application of parametric and semiparametric two-step selection models, Journal of Applied Econometrics 23(6), 729-747.
* **Kaiser, U. (2004)** Private R&D and Public R&D subsidies: Microeconometric Evidence from Denmark, Working Paper, 2004-19.
* **Klette, T. J., Moen, J. (2012)** R&D Investment Responses to R&D Subsidies: A Theoretical Analysis and a Microeconometric 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 Business Economics 24(1), 53-62.
* **Koski, H. (2008)** Public R&D Funding and Entrepreneurial Innovation, Discussion Papers 1142, The Research Institute of the Finnish Economy (ETLA).
* **Lach, S. (2002)** Do R&D Subsidies Stimulate or Displace Private R&D? Evidence from Israel, Journal of Industrial Economics 50(4), 369-90.
* **Lööf, H., Heshmati, A. (2005)** The impact of public funding on private R&D investment: new evidence from a firm level innovation study, CESIS Electronic Working Paper Series, Paper No. 06, Stockholm.
* **Meeusen, W., Janssens, W. (2001)** Substitution versus additionality: Econometric evaluation by means of micro-economic data of the efficacy and efficiency of R&D subsidies to firms in the Flemish Region, CESIT Discussion paper No 2001/01.
* **Ozcelik, E., Taymaz, E. (2008)** R&D support programs in developing countries: The Turkish experience, Research Policy 37, 258–275.
* **Reinkowski, J., Alecke, B., Mitze, T., Untiedt, G. (2010)** Effectiveness of Public R&D Subsidies in East Germany. Is it a Matter of Firm Size?, Ruhr Economic Papers 0204, Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Ruhr-Universität Bochum, Universität Dortmund, Universität Duisburg-Essen.
* **Streicher, G., Schibany, A., Gretzmacher, N. (2004)** Input additionality effects of R&D subsidies in Austria, TIP Working Paper 04-03, Vienna.
* **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.
* **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.

1. In the introductory paragraph and in Section 2 below, we refer only to the intellectual and policy context most directly relevant to the introduction and subsequent use of R&D subsidies. Hence, we make no reference to later – e.g. evolutionary, behavioural and systems – models of innovation and approaches to public intervention. [↑](#footnote-ref-1)
2. Public choice theory is the subset of positive political theory that models voters, politicians, and bureaucrats as making self-interested rational choices in an environment where outcomes are hard to measure or even define. A useful “primer” is Butler (2012; in relation to the following two paragraphs, see pp.36 and 88-94). [↑](#footnote-ref-2)
3. For discussion of the practice of cherry-picking and evidence of its adverse effects, see Radicic et al. (2014; and forthcoming). [↑](#footnote-ref-3)
4. For narrative reviews of the relevant literature see David et al. (2000), Klette et al. (2000), Cerulli (2010) and Zúñiga-Vicente et al. (2014). [↑](#footnote-ref-4)
5. Studies may report estimates that fall in more than one category of Table 1. In this case, there is one study that reports estimates under both Column 1 (12 studies) and Column 3 (4 studies). Hence, there are a total of 15 studies (not 16) in these two Columns combined. [↑](#footnote-ref-5)
6. In addition, four studies with 15 estimates identify separate additionality or partial crowding-out effects. [↑](#footnote-ref-6)
7. Appendix B discussing excluded studies is available upon request. [↑](#footnote-ref-7)
8. At the time of writing, the working version of this paper was used, published online in 2012. [↑](#footnote-ref-8)
9. One from the Input I subsample; three from Input II; and 39 from Output. In addition, 10 outliers were removed from Input I. [↑](#footnote-ref-9)
10. On funnel graphs in meta-analysis, see Stanley and Doucouliagos (2010). [↑](#footnote-ref-10)
11. The presence of only one observation smaller than zero (-0.0029) suggests an extreme reluctance to publish (or submit for publication) estimates suggesting an over-full crowding-out effect. See Stanley (2005, p.316) for a similarly extreme example. [↑](#footnote-ref-11)
12. According to the guidelines of Doucouliagos (2011), PCCs can be characterised as either “small” (PCC<0.07), “moderate” (0.07≤PCC≤0.33) or “large” (PCC>0.33). Doucouliagos (2011, p3) also addresses the question as to whether it is ‘possible to assess whether a partial correlation signifies a practically significant effect’. ‘Preliminary’ guidelines are derived from 22,000 partial correlations collected from diverse economics literatures and suggest ‘whether a partial correlation is big or small relative to what is typically found in empirical economics’ (pp.4 and 7). While recognising that the thresholds are, perforce, ‘arbitrary’, Doucouliagos (2011, p.7) judges that *small* effects are those ‘smaller than the first quartile of all partial correlations’ and are ‘of little (small) practical significance’. Of greater practical significance are the *moderate* effects between the first and the third quartiles and the *large* effects above the third quartile. [↑](#footnote-ref-12)
13. The relationship between the PCC and the t-statistic from regression *i* together with the corresponding transformation of the dependent variable is explained in Pugh et al. (2011a). [↑](#footnote-ref-13)
14. We follow Stanley and Doucouliagos (2012) and do not report random-effects estimates, since the crucial assumption of independence between the random effects and the independent variable(s) – either standard error *SEi* alone or both *SEi* and *SEi2* – is violated in the presence of publication bias. Publication bias reflects specification search and selection to offset imprecise estimates (larger standard errors) with larger estimates to obtain statistical significance. In turn, because the study-specific characteristics causing imprecise estimates – e.g. small sample size – are captured by the random effects in the composed error term, they must be correlated with the standard errors. Evidence from simulations that, in the presence of publication bias, fixed-effects estimation provides less biased estimates than does random-effects estimation is reported by Stanley (2008) and Moreno et al. (2009). [↑](#footnote-ref-14)
15. In Equation (4), the standard RESET test for the Input II subsample reveals a borderline failure of functional form at the one per cent level (p=0.007). The rejection of the null of linearity is more definite in the Output subsample (p=0.004). [↑](#footnote-ref-15)
16. 12 PCCs from the Input I and 38 from the Input II subsamples are derived from models not addressing the potential endogeneity of R&D subsidy (there are no such observations in the Output subsample). Excluding these observations does not give rise to estimates substantially different from those reported in Table 4 (albeit, the genuine effect slightly decreases). These results from this check are available upon request. [↑](#footnote-ref-16)
17. Koetse et al. (2010, p.218) note that part of the intuitive appeal of MRA lies in ‘the increase in statistical power of hypothesis testing when combining independent research results … pooling study outcomes provides a preferable estimate, i.e., an estimate with a smaller confidence interval’. [↑](#footnote-ref-17)
18. The Ramsey RESET test does not reject the assumption of a linear relationship for Equations (7) and (8) (p=0.160 and 0.091, respectively). [↑](#footnote-ref-18)
19. This general-to-specific approach is standard procedure in the meta-regression literature (Stanley and Doucouliagos, 2012, p.90). [↑](#footnote-ref-19)
20. A consistently significant and positive coefficient on a *Z* moderator indicates that – other factors remaining constant – the effect size becomes systematically more positive/less negative, albeit with different interpretations: in the Input I subsample, additionality becomes more likely; whereas in Input II and Output crowding out (full or over-full) becomes less likely. [↑](#footnote-ref-20)
21. Other *K*-variables prove significant in more than one specification and serve a control function. However, interpretation of how these might influence publication bias offers at best weak conjectures. [↑](#footnote-ref-21)
22. Econometric estimation is possible only for influences that vary between observations. [↑](#footnote-ref-22)
23. The values predicted by MRA models are free of publication bias. Hence, one possible explanation of failure to predict would be the presence of selection bias in the reported out-of-sample effects. [↑](#footnote-ref-23)
24. Quite apart from the date of publication, we could not include these studies in our MRA database, because neither reports the standard errors, t-statistics or, at least, the exact p-values required for precisely computing our effect size, the PCC. However, we can identify the minimum value that the t-statistics can take corresponding to the reported information (either significance at the one per cent level or that p=0.000) and to the degrees of freedom. Accordingly, we can calculate the minimum corresponding values of the PCCs for these two studies. We regard this as acceptable for the purposes of the check reported in this section. [↑](#footnote-ref-24)
25. The results reported in Table 9 are not much changed by assuming larger t-statistics. [↑](#footnote-ref-25)
26. To take an example from a different area of public policy, albeit one contributing to human capital formation and thus complementary to R&D investment in models of endogenous growth, recent studies on per pupil spending by English and Australian secondary schools report positive but very small attainment effects (Pugh et al., 2011b; Pugh et al., 2014). [↑](#footnote-ref-26)