WPS6532

Policy Research Working Paper 6532

The Impact of Government Support on Firm R&D Investments

A Meta-Analysis

Paulo Correa
Luis Andrés
Christian Borja-Vega

The World Bank
Entrepreneurship and Innovation Unit
South Asia Sustainable Development Department
Water and Sanitation Program
July 2013

Policy Research Working Paper 6532

Abstract

This paper applies meta-analysis techniques to a sample of 37 studies published during 2004–2011. These papers assess the impact of direct subsidies on business research and development. The results show that the effect of public investment on research and development is predominantly positive and significant. Furthermore, public funds do not crowd out but incentivize firms to revert funds into research and development. The coefficient of additionality impacts on research and development ranges from 0.166 to 0.252, with reasonable confidence intervals at the 95 percent level. The results are highly sensitive to the method used. The high heterogeneity of precision is explained by the wide variety of methodologies used to estimate the impacts and paper characteristics.

This paper is joint product of the Entrepreneurship and Innovation Unit, South Asia Sustainable Development Department, and Water and Sanitation Program. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at pcorrea@worldbank.org, landres@worldbank.org, and cborjavega@worldbank.org.

The Impact of Government Support on Firm R&D Investments: A Meta-Analysis

Paulo Correa, Luis Andrés, and Christian Borja-Vega[1]

Key words: Research and Innovation, Impact, Meta-Analysis.
JEL: O30, O38, D22
Sector Board: Finance and Private Sector

[1] Paulo Correa is Lead Economist at the Entrepreneurship and Innovation unit of the World Bank, Luis Andres is Lead Economist in the Sustainable Development Department for the South Asia Region of the World Bank, and Christian Borja-Vega is an Economist at the Water and Sanitation Program of the World Bank. They are grateful to Esperanza Lasagabaster and Hari Subhash for their comments and suggestions. The findings, interpretations and conclusions expressed herein do not necessarily reflect the views of the Board of the Executive Directors of the World Bank or the governments they represent. Senior authorship for this paper is not assigned. The authors' email addresses are: pcorrea@worldbank.org, landres@worldbank.org, and cborjavega@worldbank.org.

1. Introduction

The promotion of investments in research and development (R&D) and innovation is a standard component of ‘stimulus packages’ adopted by advanced economies to counterbalance the effects of the recent global crisis (OECD 2012). For example, according to Eurostat, government budget appropriations or outlays for research and development (GBAORD) increased 46 percent in the Slovak Republic, 33 percent in Korea and 20 percent in Germany in the period 2007-11 (Eurostat, 2011).

Governments have been particularly concerned with the possible decline in R&D investments by the private sector and its impact on innovation and productivity. Approximately three-quarters of OECD economies adopted new measures to foster business investments in R&D to counter this - including higher tax-credits, additional direct support or both, as in the cases of France, Japan, Norway, and the U.S. (OECD, 2011). Direct support to business R&D corresponds to about 1.27 percent of GDP on average, and overall spending in R&D (including the government and higher education sectors) reached 2.06 percent.[2]

This consensus among policy-makers, however, needs to be supported by empirical evidence to substantiate the causal links that are assumed when they allocate public funds to private R&D projects, more specifically the stage multiplier effects of R&D subsidies on R&D expenditures, especially for input, output, and outcome additionality. This literature also needs to be supplemented with a systematic review that aggregates findings to offer policy directions.

One first attempt to collect and review empirical literature on the impact of direct public support to private investment in R&D was performed by David et al. (2000). The majority of studies surveyed in this paper point out the following conclusions: Government R&D and tax

[2] Business investments in R&D were considered to be pro-cyclical due to the reduction in firm’s cash flows or simply worsening of financial market conditions. Internal funds are the preferential source of financing for R&D and innovation investments. With the global downturn, firm’s revenues declined substantively. Also, financial market conditions (cost and availability of capital) worsened significantly, reducing the availability of internal and external funds for research and innovation. For example, using a French firm-level panel data set over the period 1993-2004, Aghion et al. (2008) show that the share of R&D investment over total investment is countercyclical without credit constraints, but it becomes pro-cyclical as firms face tighter credit constraints. Other studies have argued that business R&D is pro-cyclical even when firms are not financially constrained (Barlevy, 2007).

incentives stimulate private R&D investments. Government grants and contracts, and government spending on basic research do not displace private R&D funding except when R&D inputs have inelastic supply. The outcome depends on market demand and supply conditions, which are unobserved most of the time. About two-thirds of studies surveyed by David et al. (2000) conclude that public funding is complementary to private financing, while one-third point to a substitution between the two sources.

However, there was a high degree of heterogeneity in the surveyed studies (conducted over the previous three decades), in terms of the data used, the level of analysis (micro/macro; industry/firm) and the econometric strategies. In addition, most of these studies were subject to a potential selection bias and other serious methodological limitations that disqualify the predominantly positive evidence and poses more questions than concrete answers about the relationship between private investment and government support to R&D. The relationship between the two also depends on the level of aggregation of reported studies and on the country studied. Studies based on a lower level of aggregation (line of business and firm data) tend to report substitution almost as often as 'complementarity' [47% of all studies, and 58% of US studies respectively (the US studies represent two-thirds of all surveyed studies)]. The authors note that the tendency of aggregate studies showing a complementary relationship could be result of: i) Positive covariation of public and private components and inter-industry differences in technological opportunity; and/or ii) The effect of government funding of R&D raising the cost of R&D inputs to private R&D activity.

This paper contrasts and combines the results from 37 papers published during the 2004-11 period in order to identify patterns among study results, sources of disagreement among their results, or other interesting relationships that may come to light in the context of R&D interventions. We used Meta-Analysis techniques that aim to combine studies with similar research questions to increase precision and assess the generalizability of results. The regressions ran this precision value against the standardized degree of correlation between papers. The precision variable contains the standardized measure of the impact estimators of R&D additionality reported in each paper. This systematic review aims to obtain a better understanding of the impact of these interventions. The analysis suggests that based on the surveyed studies there is indication of positive R&D impacts. Furthermore, public funds do not crowd out but incentivize firms to revert funds into R&D. Results show that the effect of public investment in R&D is predominantly positive and significant. The coefficient of additionality impacts on R&D ranges from 0.166 to 0.252, with reasonable confidence intervals at the 95 percent level.

However, the estimation results have a large range, suggesting that the meta-data is highly sensitive to the method used. The small sample size of surveyed studies produces high ranges of confidence intervals across independent variables coefficients across different models. The high heterogeneity of precision is explained by the wide variety of methodologies used to estimate impacts. The use of "gold standard" evaluation methods (randomized assignment) is not common at all in the literature sample. Hence, the findings of this paper should be reconfirmed with more rigorous impact evaluations techniques and periodic updates with subsequent studies that could increase the sample size of the meta-analysis exercise.

The paper is organized as follows: Section 2 summarizes the existing literature regarding impacts in R&D in programs while Section 3, describes the data and methodology used in this paper. It also presents the results of the methodology. Section 4 summarizes the main findings.

2. Literature Review

Since the last meta-analysis (David et al. 2000) there have been several studies that tried to correct biases and methodological weaknesses in earlier studies using larger firm-level panel data, quasi-experiments, and better econometric techniques (such as propensity score matching). These studies have brought greater homogeneity to the literature, and improved the quality of the analysis. Given this new and more robust evidence, it is important to conduct a systematic review of the literature once more, particularly because of the current emphasis on increased government spending on R&D. However, to the best of our knowledge, no such review of these newer studies on the impact of direct public support to R&D has been undertaken yet. This paper uses meta-analysis techniques to systematically review this new body of literature. From a larger sample of papers, we identified a 37 studies published during the 2004-11 period in which different techniques are applied to increase robustness of results.

In this section, we present a survey of the newer literature, highlighting the key results and the methods used to address some of the biases and methodological flaws in studies before year 2000. Our survey of literature shows that while impact estimates (for different aspects of R&D interventions) tend to shrink compared to earlier, less rigorous studies, the effect of public investment in R&D still remains predominantly positive and significant.

Some of the sources of overestimations in studies about the impact of public support to business R&D can be explained by:

- **Specification and Endogeneity**: This factor is distinguished by the application of models from a structural- to non-structural-analytical perspective. The former implies that the outcome equation and the selection-into-program are separately modeled in a system of simultaneous equations, and encompasses macro or aggregated outcomes. The latter implies only the inclusion of outcomes for the purposes of analyzing specific sectors or types of firms.

- **Data**: Models are based on a cross-section dataset, pooled data and/or longitudinal datasets (allowing for dynamic and long-run analysis). Very few studies actually collect data for the purpose of answering specific questions about R&D impacts. The majority use existing survey data and/or administrative records.

- **Policy Variables Assessed**: Models using a binary policy variable (generally in the form of "subsidized" versus "non-subsidized" units), and models using the policy variable in levels (i.e., in a continuous form) have shown to be irrelevant as it is important to build a statistically robust counterfactual to make valid comparisons between firms. In addition, there are several unobserved attributes that cannot be separated from this variable, which may pose mixed or confounding conclusions.

- **Identification Strategy**: Papers that robustly evaluate the impacts of R&D by addressing the issue of causality. The majority of the studies used matching techniques, instrumental variables or sample selection correction as part of the non-experimental retrospective approach to evaluate. Very few studies have used regression discontinuity to identify impacts by utilizing a cutoff criterion to separate comparison groups. A handful of studies use micro-simulation approaches, and there are very few studies that actually identify impacts through a randomized trial.

The main issue in these studies is the large bias, the most common being selection bias, i.e. the so-called 'treatment' group (e.g. recipients of public funding) since the interventions covered in those studies were usually not implemented in a random fashion. Instead, in general, governments cherry-pick projects with the highest expected (social) value. Several recent studies attempt to handle this bias using matching methods.

For instance, Almus and Czarnitzki (2003) use matching methods to find an overall positive and significant effect of R&D subsidies on investment in R&D by firms in Eastern Germany. Gonzalez et al. (2005), estimate the probability of obtaining a subsidy, assuming a set of firm's observables as pre-determined (e.g. size, age, industry, location, capital growth), to identify a very small but positive effect of R&D grants on private investment (significantly larger for small firms) in Spain. Gorg and Strobl (2007) combine the matching method with Difference-in-Differences (DID) estimation to find that in Ireland small grants had additional effects on private R&D investment, while large grants crowded out private investment. Lopes Bento (2011) studies the effect of public funding on internal R&D investment and on total innovation intensity at a cross-country comparative level. Applying a nonparametric matching method to identify the treatment effect, the author finds that on average firms would have invested significantly less if they would not have received subsidies. While, Hussinger (2008) uses two-step selection models to show that German public subsidies were effective in promoting firms' R&D investment. On the other hand, robust studies that used instrumental variables or simple comparisons between subsidized and non-subsidized firms (Lach, 2002; Wallsten, 2000) find null effects on innovation intensity, output productivity and innovation efficiency.

Lopez et al. (2010) who study the effects of the Argentinean Technological Fund (FONTAR), test the additionality versus crowding-out hypothesis, i.e. evaluating whether the presence of the public aid to innovation complements or crowds out a firm's investments in innovation activities by modeling the impact of all FONTAR's programs and of the ANR (Non-technological subsidies) program on total and private innovation expenditures. They find that beneficiary firms spend more on innovation activities (e.g. research and technology purchases), even when the amount subsidized or granted is netted out from the total amount spent (asymmetry correction). Marino et al. (2010), however, find slightly contrary evidence. They use a continuous treatment evaluation design to identify the marginal effects of treatment and use it to determine sub-optimal amounts of funding. Their results indicate a high level of substitution between public and private funds at higher levels of government subsidies.

Baghana (2010) on the other hand, uses a conditional semi-parametric difference-in-differences estimator on longitudinal data to analyze the impacts of public R&D grants on private R&D investments and on the productivity growth of the manufacturing firms in a context where fiscal incentives (like tax-credits) are present. The results show that the effect of fiscal incentives on firm productivity and input additionality is enhanced when combined with subsidies. These results show that the choice for policymakers is less between fiscal incentives and subsidies rather it is more about identifying the suitable level of additional funding (subsidies and grants) when fiscal incentives like tax-credits are already provided.

The papers reviewed so far highlight the effects of increased R&D spending by governments. The link between R&D spending and innovation however, might not work when these are structural flaws in the R&D system as highlighted by Roper (2010) in a study of Western Balkan countries. Based on an econometric examination of the innovation production function he shows that increased R&D spending and skill development does not lead to an increase in innovation due to structural flaws. He therefore suggests an active and rather interventionist innovation policy in the Western Balkans countries to address these system failures.[3] The limited impacts vary when the institutional setting is included in the model. This suggests that observed features of innovation systems may also be posing R&D impact biases. Based on an econometric examination of the innovation production function in each area, they observe marked differences in the determinants of innovation. The author finds that R&D program biases are not strongly related to a firm's characteristics but rather on the program's institutional setting.

[3] Interventionist policies influence the pace of R&D in private firms. Active policy interventions refer to a "hands off" approach where businesses have more flexibility to decide where to allocate their subsidized resources. These types of interventions are recommended in the presence of institutional and R&D system improving performance. Only small pilots have addressed the active R&D policies in the Western Balkans countries.

Hall and Maffiolli (2008) combined administrative records in Argentina, Brazil, Chile and Panama with innovation and industrial surveys and use quasi-experimental (matching) methods to test four aspects of R&D impacts: input additionality, behavioral additionality, innovative outputs, and performance. After correcting for selection biases in all countries (using techniques such as propensity score matching, difference in differences estimation, fixed effect panel estimation, and instrumental variable estimation) the authors found positive and significant effects on input (intensity) and behavioral (firm proactiveness) additionality. While innovative output impacts were found to be positive but statistically insignificant after bias correction (due to smaller sample size in this area). Finally, in terms of a firm's performance, positive impacts were found on firm growth but not on its productivity. This research consistently highlights the tendency to overestimate impacts in the absence of a statistically valid counterfactual, regardless which aspect of the firm is influenced by R&D subsidies.

Sayek (2009) separates the effects of R&D stimulus according to performance and FDI financing. The author debates the effectiveness of (or additionality of) public R&D spending and the productivity impact of private sector R&D spending. She finds that the relation between the government R&D activity and private sector R&D activity seems to be stronger in financing than performance.

Conversely, Czarnitski et al. (2004) explores a similar separation effect (performance and financing) but with tax credits as the main intervention to foster R&D. The authors find that not only do tax credits have a positive impact on a firm's decision to conduct R&D investments but also on higher product innovations. Other findings include: i) Fiscal incentives have a short run effect on private R&D, whereas government R&D is stimulating in both the short and long term; ii) The size of the impact of R&D subsidies varies with respect to the subsidization rate and has an inverted-U shape, denoting increasing effectiveness associated to government R&D up to a threshold that ranges from 5 to 25 percent and decreasing effectiveness beyond; iii) The more stable the policy instruments, the more efficient they are in stimulating private R&D; and iv) Policy tools and incentives in R&D appear to be substitutes, raising one of them reduces the stimulating effect of the other.

We have gained some important points from surveying recent impact evaluation literature on the effects of R&D interventions on firm-specific outcomes. Conceptually, the majority of R&D evaluations are still characterized by having weak evaluation designs. It is also clear from the literature review that the better the evaluation method used, the more statistically robust the effects, leading to unambiguous conclusions regardless the area of R&D explored. In addition, surveyed evaluations focus on different aspects of R&D impacts, but in general such impacts tend to shrink due to biases. Certain research areas of R&D still have important knowledge gaps because of lack of data and rigorous evaluation designs. Finally, not only do the scale of intervention, size of the firm, type of beneficiary, project attributes, and other variables contribute to explain biases, but other study-design variables can contribute to explain impact biases.

3. Meta-Analysis of R&D Impact Studies

With the mixing results from the literature, a meta-analysis of recent papers can help to verify if the claims about the existence of weak methodologies resulting in biased impact estimates are true. Meta-Analyses aim to combine studies with similar research questions to increase precision and assess the generalizability of results. The regressions ran this precision value against the standardized degree of correlation between papers. The precision variable contains the standardized measure of the impact estimators of R&D additionality reported in each paper. The precision coefficient indicates if there are effects covered by papers in the sample: if this coefficient is equal or close to zero the papers have low precision in their estimators. All studies selected for the meta-analysis were selected on certain evaluation characteristics and R&D themes. This section describes the steps undertaken in the meta-analysis.

Extensive searches were run in order to identify the order of magnitude of papers to be included in the database.[4] The searches were initially carried out with three search engines:

[4] A comprehensive search was carried out to identify all econometric evaluations reporting estimates of R&D on a firm's outputs and outcomes. Numerous keywords were used for the search process. The search carefully checked references cited within empirical, theoretical, and review studies. Both published (books, reports and journals) and unpublished (working papers, dissertations) studies were searched only in English. The search covered a duration of three months and ended in June 2011.

Google Scholar, JStor, and Elsevier.[5] The variables collected from each study covered the methodological approach, the estimation strategy, the type and characteristics of the interventions, and the overall impact estimates.[6] The studies considered in the analysis needed to show a formal evaluation methodology applied to an existing R&D intervention. Despite the fact that the paper search covered the period from 2004-2011, most of the papers with formal evaluation methodologies are recent (2008-2011). Thus, a database with 37 papers was constructed (Annex 3 includes the list of papers covered for this database).[7]

Once the database was set up, Meta-Analysis Estimations (MAE) were conducted. MAE methods have advanced enormously in the past five years in terms of analytical procedures with more precise estimations. Not only can MAEs depict plots and graphs with the average effects of the papers collected for conducting Meta-Analysis, but they can also show the degree of heterogeneity, precision and bias of estimates from each publication (Doucouliagos and Stanley, 2008). In addition, several descriptive statistics procedures in MAEs shed light onto the quality of the paper data collected and the feasibility of conducting unbiased Meta-Analysis Regressions (MARs).

It is important to highlight that the use of MAEs and MAR are useful even when dealing with a small sample of papers from a larger literature branch. This is because although systematic reviews and meta-analysis have the potential to produce precise estimates of treatment effects that reflect all the relevant literature from a particular topic, they are not immune to biases.

One important point to consider is how the 'additionality concept' differs between studies. However, as the concept of behavioral additionality is quite flexible, evaluators may take into consideration a range of behavioral changes. One of the major shortcomings of a meta-analysis from R&D evaluations is that important qualitative attributes of the programs are not assessed. Behavioral additionality cannot capture the strategic relevance of funded projects as it is

---

[5] Although the engine search delivered more than 120 publications only 37 were included in the meta-analysis because they meet the criteria of a quantitative evaluation. In other words, these publications used data and a formal methodology to estimate the effects of R&D programs, beyond solely reporting simple correlations. It was important to only include studies with impact estimates (and their standard errors) since they are relevant when undertaking meta-data analysis.

[6] See Annex II for the complete list of variables used and their categories.

[7] Although 40 papers are listed in the Annex, 37 were included because 3 papers did not have complete information to be part of the analysis.

perceived by the beneficiaries. The fact that beneficiaries are encouraged by the policy to do something that they do not perceive as strategically relevant could be a positive result of the intervention assuming that policy-makers have a clearer and better understanding of the future perspectives and evolutions (Lukkonen, 2000).

3.1 Data and Methodology

The data used for this analysis was collected from papers that explicitly targeted evaluating R&D programs quantitatively.[8] Four categories of variables were collected from the papers and from the R&D programs. The first category of variables relate to the methodology and the estimation method used for the evaluation (which includes the type of data), and evaluation strategy used to identify impacts. The second category of variables relates to a paper's attributes, including year and type of publication, region/country covered, and literature gap addressed among others. The third category of variables relates to R&D program attributes, containing the funds available, number of firms covered, types of institutions managing, and granting funds. Finally, the fourth category of variables relate to the quality of the publication proxy by citations, web references, and paper access statistics.

Differences among studies may be categorized broadly into those related to the phenomenon being studied and those unrelated. Choice of study design may induce differential biases in the results as well. As mentioned above, simple Ordinary Least Squares (OLS) will produce biased and inefficient estimators due to the implicit heterogeneity across papers and estimations. A meta-regression (MAE/MAR) can be either a linear or logistic regression model. In most meta-regression approaches, the unit of analysis, that is each observation in the regression model, is a study. But studies differ in quality. This may lead to publication bias because poor quality studies were not selected as surveyed studies. However, there are methods to correct for publication biases using MARs -by allocating weights depending on the relevance and quality of the publication (see figures B1-B4 in Annex 1 for publication biases in our database), among other attributes. The questions that a meta-analyst may answer with a meta-regression include

[8] Two qualitative studies that used survey data from firms were included.

estimating the treatment effect controlling for differences across studies, and determining which study-level covariates account for heterogeneity.

The method used for the meta-analysis (MAEs/MARs) aims to combine all comparable estimates from different studies and to draw inferences from these with respect to: i) the existence of horizontal and vertical variability (Figures 1 and 2); ii) the size of the interactions between papers; and iii) the factors that explain the wide variation in reported estimates. The MRA model involves regressing comparable measures of an effect (partial correlations) against a constant and a set of variables that can explain the heterogeneity in estimates, such as data, specification and estimation differences in research design:

$$r_{ij} = \beta_0 + \beta Z + v_{ij}$$

Where r_{ij} are partial correlations, for the i^{th} estimation (impacts), from study j. Z_{jk} are moderator variables used to explain the large within and between study heterogeneity routinely found in economics research (Stanley and Jarrell, 1989). The vector Z in our case contains information on the type of data used for the analysis in each paper, region or country covered by the paper, and some qualitative aspects of the papers like estimation method used and paper precision in terms of significance of relevant impact estimators, sample size, etc. Finally, v_{ij} is the random error term.

It is important that in the context of these regressions, some descriptive meta-analysis statistics are presented to verify if the papers covered by the review have explanatory power. Publication bias is frequently found in meta-analysis—the association of publication probability with the statistical significance of results (Stern and Simes, 1997). In addition, this meta-analysis of R&D evaluations requires the verification of the existence of biases and heterogeneity in each study's results. Because many used studies used counterfactuals, the treatment impacts also need to be assessed statistically to prove that the sample of the meta-analysis can draw conclusions from the literature.

Heterogeneity may arise from genuine empirical differences in the underlying R&D models and functions, but it can also arise from misspecification of the econometric models. MRA helps to quantify both the effects of misspecification and the genuine differences in strategic interactions. Given that the majority of the evaluation studies surveyed have impact estimators and their corresponding standard errors, the partial correlation coefficient between studies take the following form:

$$r_{ij} = (\beta_{ij}) * \left[t/\sqrt{t^2 + df)}\right]$$

where β is the impact coefficient, t is the t-value of estimation and *df* are the degrees of freedom resulting from the estimation. When analyzing how the estimates between each other compare it is necessary to take into account the precision of the estimates, which is approximated by the inverse of the standard error of the impact estimations (Costa-Font et al., 2011).

The method considered to estimate the meta-impacts from all studies is not limited to simple OLS. This is because estimates reported between studies might not be statistically independent of each other, which violates one of the OLS assumptions. Therefore, to solve this problem, truncated and meta-regressions were estimated to approximate restricting estimates based on clustered publications. The mean R&D effect is the weighted average of the standardized effects derived from each study (e.g, simple correlation, partial correlation or elasticity between R&D investments and firms' outcomes). It is customary to use a weighted mean, ε, because studies differ in the amount of information they offer. Although we also experiment with the Impact Factor of the journals in which the studies are published, it is a standard practice in meta-analysis to the use sample size as the weight. Partial correlations measure the impact of R&D on firms' performance holding other factors constant.

In order to estimate weights for the regressions, the precision of each paper was estimated in order to compute a Weighted Least Squares (WLS) where papers with lower standard errors reported and higher sample sizes get higher weights. One important aspect that is included in the meta-analysis deals with the evaluation method used and the type of data utilized in each paper. Dummy variables were created for each category of evaluation method and data type used, regressed against the t-values of the impact estimators and the meta-regressions. Controlling for these factors improves the specification, data and methodological differences on the results of the studies (Doucouliagos and Ulubasoglu, 2010).

Theory, informal impressions, and anecdotal evidence suggest that the estimated R&D impacts to research are likely to have been affected by decisions made by analysts about the specification of the models, which as a consequence might bring biases to the meta-estimators. However, the evidence shown in this paper suggests that some publication biases might exist depending on the publication's quality and program characteristics. The meta–analysis also intends to shed light on the variation of the partial correlation coefficient between papers affected by the estimation model specification, such as sample choice, type of estimator, inclusion/exclusion of control, paper and program variables for R&D impacts.

The strategy to identify the average impacts of the evaluation studies consists of four stages. First, simple OLS regressions were run to test the signs and significance levels from publication and method variables against the t-values of the estimators and the partial paper correlations between studies. In the presence of any type of bias OLS estimates will not produce credible average impacts from the studies surveyed. Because of this issue, the second stage consists of exploring the reliability and heterogeneity of the meta-data by plotting meta-graphs, such as funnel plots.[9] After this step, meta-regressions are run to i) identify the sources of bias, and ii) estimate the weights for each publication. With these weights, the final stage consists of estimating a WLS that ponders low weights in the papers cause a higher bias and a higher weight to those with relatively stronger methodologies.

Funnel graphs are the conventional methods used to identify publication selection. A funnel graph is a scatter diagram of precision (1/standard error) versus estimated effect. Funnel graphs can also plot estimation errors between-paper correlation, where one would expect that as papers are more comparable they will line up at zero. In the absence of publication selection, the diagram should resemble an inverted funnel. Asymmetry is the mark of publication bias. To corroborate this pictographic identification of publication bias, we use a meta-regression

[9] A funnel plot is a simple scatterplot of intervention effect estimates from individual studies against some measure of each study's size or precision (Light and Pillemer, 1984; Begg and Berlin, 1988; Sterne and Egger, 2001). It is common to plot effect estimates on the horizontal axis and the measure of study size on the vertical axis. This is the opposite of the usual convention for two-way plots, in which the outcome (e.g., intervention effect) is plotted on the vertical axis and the covariate (e.g., study size) is plotted on the horizontal axis. The name "funnel plot" arises from the fact that precision of the estimated intervention effect increases as the size of the study increases. Effect estimates from small studies will therefore scatter widely at the bottom of the graph, with the spread narrowing among larger studies. In the absence of bias, the plot should approximately resemble a symmetrical (inverted) funnel. Funnel plots are commonly used to assess evidence that the studies included in a meta-analysis are affected by publication bias. If smaller studies without statistically significant effects remain unpublished, this can lead to an asymmetrical appearance of the funnel plot.

analysis (MRA) of the t-value versus precision (Egger et al., 1997). This follows the specification as:

$$\mathrm{effect}_i = \beta_1 + \beta_0 \mathrm{Se}_i + e_i$$

The reasoning behind this model of publication selection begins with the recognition that researchers will be forced to select larger effects when the standard error is also large. Large studies with smaller standard errors will not need to search as hard or long for the required significant effect. Accounting for likely heteroskedasticity leads to the WLS version of the following equation:

$$t_i = \beta_0 + \beta_1 \left({}^{1}/_{\mathrm{Se}_i}\right) + e_i$$

In the absence of publication selection, β_0 will be zero and the precision estimated coefficient adjusted through WLS run against the between study correlation will also tend to zero. Without selection, the magnitude of the reported effect will be independent of its standard error. Precision's regression coefficient also serves as a test of genuine empirical effect beyond publication bias. As suggested by Monte Carlo simulations (Stanley 2004), it is prudent to confirm this positive precision-effect test with another MRA test for genuine effect. However, due to the reduced sample of our database the simulations cannot be computed. In sum, the MRA method tests evidence of an authentic effect and the average effect also needs to be statistically significant.

Extracting multiple effect sizes (see Annex 2 for detailed method) from a single study, however, might result in a violation of the independent assumption for effect sizes, which in turn, might increase Type I or II errors (Glass et al., 1981). In this study, two approaches were employed to resolve this dependence problem. First, only one finding per outcome was extracted from each study unless they represented different programs.[10] This approach enabled us to examine different outcomes while ensuring independence among the findings for each outcome. Secondly, multiple effect sizes provided by the same program for the same category of outcome were dealt with by randomly taking a single value from the set of correlated effect sizes per

[10] This was just the case for two studies.

feature for each affected study. This method eliminated the problem of dependency while ensuring that all levels of a study's features were represented (Lou et al., 2001).

Meta-analysis focuses on the direction and magnitude of the effects across studies, not statistical significance. In other words, the meta-analysis mean impact coefficient will show the direction of impacts, and whether this direction is statistically significant after correcting for sources of selection and heterogeneity. In addition, there are some aspects worth highlighting from this specific meta-analysis. First and foremost, the collection of research data is entirely empirically based on concrete evaluations. Second, the meta-analysis produces quantitative results on the direction of impacts with limited consideration for other qualitative aspects of the studies. Third, the findings can be configured in a comparable statistical form (e.g. effect sizes, correlation coefficients, odds-ratios, proportions). Finally, it is also worth highlighting that despite publication selection, biases may still pose limitations because negative and null finding studies may have not been published.

In sum, the advantages of meta-analysis (e.g. over classical literature reviews, simple overall means of effect sizes etc.) include: i) Derivation and statistical testing of overall factors/effect size parameters in related studies; ii) Generalization to the population of the studies; iii) Ability to control for between-study variation; iv) Including moderators to explain variation; and v) Higher statistical power to detect an effect than in 'n=1 sized study sample.

Conversely, the review included studies that mainly follow ad hoc identification strategies for the evaluation. Prospectively randomized evaluation designs in the R&D literature are few. Most evaluations tend to adapt their methodologies to existing data. This poses a weakness in the methodology because a good meta-analysis of badly designed studies may still result in statistics with little reliability. However, given that all the meta-analysis in this paper relies on published articles this weakness might be substantially reduced.

3.2 Descriptive Statistics and Graphs

Descriptive statistics have two objectives. First, they show the distribution and composition of variables in our dataset. Second, they depict the sources of biases across publications and the direction of such biases. Table 1 shows the means of some relevant variables from the Literature Review database used in this analysis. More than half of the papers included in the review used surveys to evaluate R&D additionality, and one out of six papers used administrative records for the same purpose, while close to 1 out of 5 used more than one source of data to assess impacts. With regards to the type of data arrangement used for the evaluation, around 51 percent of surveyed papers used a panel structure, and 27 percent used pooled cross section (or time series) for the analysis. Overall, many papers used panel data structure in order to quantify for contemporary or time-dependent factors that may play a partial role in determining the magnitude and sign of the impacts. The estimation methods used to assess impacts relied mostly on quasi-experimental approaches (35 percent of all studies), including propensity score matching and regression discontinuity, and around 11 percent of the studies relied on micro-simulation techniques to build a statistically reliable counterfactual. More than 40 percent of studies used structural models or instrumental variables to assess impacts, which are indicative of the small proportion of papers left that prospectively designed and evaluate R&D programs by using randomized assignment into the comparison groups.

Table 1. Database Summary Statistics (IE)

	Data Used for Analysis	Type of Data	Methodology
Survey	56.76		
Admin. Records	16.22		
Mixed	18.92		
No Data Used	5.41		
Not Specified	2.7		
Panel		51.35	
Cross-section		16.22	
Pooled Cross Section		27.03	
Not Specified		5.41	
Quasi-experimental			35.14
Simulation			10.81
Structural model/IV			43.24
Qualitative/other			10.81

The main characteristics of each publication and the R&D programs evaluated were also collected. Table 2 shows that the average year of publication is 2010 and around 70 percent of the publications explicitly developed methodology that used a counterfactual to assess impacts. The average sample size of the data used to conduct the assessment in each paper is 2,143 observations with a relatively high standard deviation.[11] Around 70 percent of R&D programs included in the papers were preceded by laws and enactments and only 43 percent were in the later stages of implementation. On average the impact coefficient was 0.32 with a standard deviation of 0.13. The average cumulative investment of R&D programs surveyed was 175 USD million (PPP).

On average, each paper had around 110 web views, taken from each website/publication traffic statistics. But only some of these papers were cited in other publications 15 times on average. Around 80 percent of papers and publications followed a peer review process. Only 1 out of 4 papers had sections devoted to robustness checks of impact estimates.

Table 2. Database Summary Statistics (Program and Papers)

	Mean	*Std Dev.*	*Min*	*Max*	*N*
Program Characteristics					
Year of Publication	2010	2008	2004	2011	37
Use of Counterfactual	0.70	0.46	0.00	1	37
Sample size	2,143	2,968	66	12,566	31
Firms Benefited by Program	103,305	170,236	1,000	761,000	34
Late Stage of implementation	0.43	0.50	0.00	1	37
Program preceded by Law	0.70	0.46	0.00	1	37
Years of Implementation	12.49	7.89	4.00	38	37
Impact Estimator	0.32	0.30	0.03	1	33
Standard error estimator	0.13	0.14	0.01	1	33
Total cumulative program investment (PPP, usd millions)	175	324	3	1,500	35
Paper Characteristics					
Number of Views (web-based)	111.41	115.83	11.00	552	37
Citations (google_scholar)	14.59	23.31	0.00	104	37
With Fast View Server	0.54	0.51	0.00	1	37
Number of pages for analysis	10.23	6.84	3.00	39	37
Paper peer review	0.81	0.40	0.00	1	37
Rate of paper quality (=10 best)	7.27	1.16	5.00	10	37

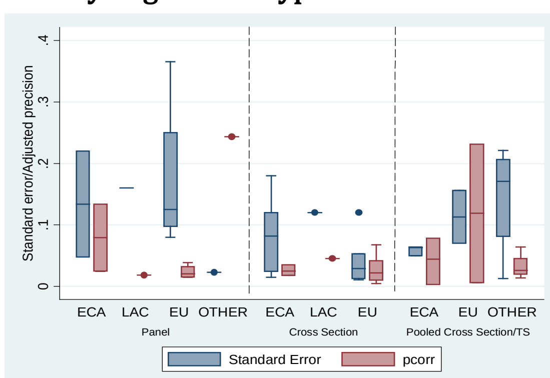
[11] From the 37 papers surveyed the smallest sample sized used was 66 and the largest was 12,500.

Paper with Section Robustness checks	0.24	0.43	0.00	1	37

Source: Own Estimation based on R&D Evaluations Review

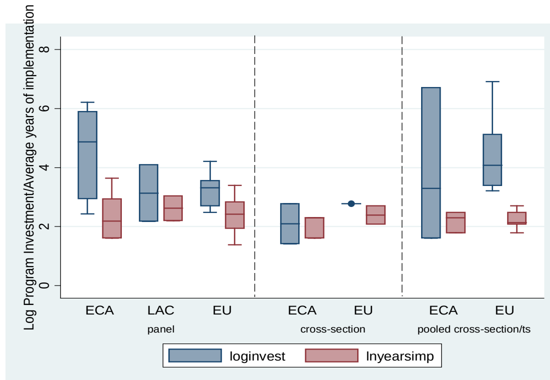
Figure 1 shows the heterogeneity in the papers included in the meta-analysis. The papers are separated by Region and Type of data used, plotted against the precision of estimates and then by R&D program characteristics (years of implementation and investment). In both cases, there is wide variability, particularly in papers that evaluate R&D projects in European Union (EU) and used panel data (Figures 1 and 2). The meta-comparability of papers depicts comparability across publications given their number of citations (Figure 3). The size of the circles indicates the number of citations and the closer and more overlapped the circles are, the higher their comparability. In general terms, given the small size of the sample there is sufficient comparability from a graphical perspective.

Figure 1. Variability in Papers' Standard Errors and Correlations (Between Estimations) by Region and Type of Data Used



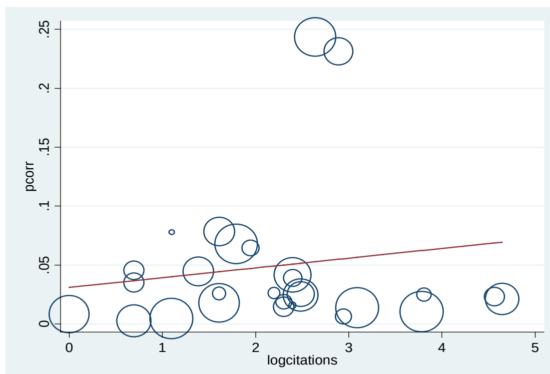
Source: Own estimations

Figure 2. Variability in Papers' Cumulative Program Investment and Years of Implementation by Region and Type of Data Used



Source: Own estimations

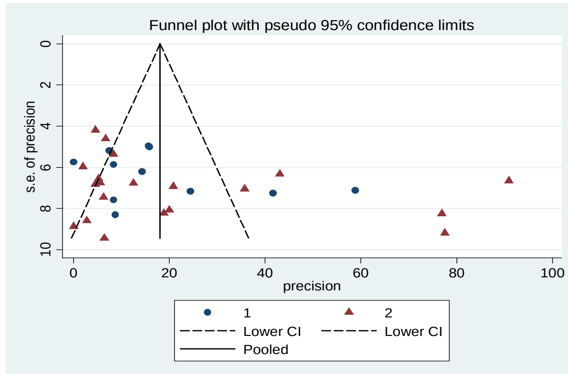
Figure 3. Number of Citations and Between Paper Correlations (Publication Weight)



Source: Own estimations

If evidence of heterogeneity in the effect of treatment between studies is found, then meta-regression can be used to analyze associations between treatment effect and study characteristics. This is one of the main reasons to incorporate characteristics such as peer reviewing, downloads, citations, pages, and so on. The descriptive graphs impose a benchmark on the process of summing up research findings. It also helps represent findings in a more differentiated and sophisticated manner than conventional statistics. For instance they can show depictions of relationships across studies that are obscured in summary statistics and protects against over-interpreting differences across studies. They can also handle a large numbers of studies (this would overwhelm traditional approaches to review). Figure 4 shows papers' comparability fit given their precision and standard errors of their precision with a single criterion for all pooled papers. In a similar fashion figure A1 (Annex 1) shows precision and comparability of those studies with similar methods used to identify impacts.

Figure 4. Most Papers Fit Precision Comparability



1=Precision, 2=Log Sample Size

Source: Own estimations

Table 3 shows the results for the heterogeneity tests. There is large heterogeneity across studies which rely on quasi-experimental and structural models-98 and 96 percent respectively, significant at 99 percent level. The simulation and qualitative approaches to evaluate R&D interventions show null heterogeneity. This has to do with the standardized procedures to collect qualitative data and the few alternative methodological approaches that authors have to conduct simulations.

Table 3. Heterogeneity Test

Tests for Heterogeneity Between Evaluation Method Groups				
Type of Evaluation Model	**Heterogeneity Statistic**	**P-value**		**I-squared ***
Structural model	1041.1	0.000		98.6%
Quasiexperimental	236.9	0.000		94.9%
Qualitative	3.0	0.394		0.0%
Simulation	1.9	0.588		0.0%
Overall	1731.5	0.000		97.9%
** Variation in ES attributable to heterogeneity*				

3.3 Main Results

Table 4 shows the simple OLS regressions for three different specifications.[12] The OLS regressions register r-square of around 0.40 indicating acceptable fit of the variables explaining the outcomes and, in particular, both OLS models show positive and significant signs for the precision variable. Costa-Font et al. (2011a) explain that when this parameter is different than zero, then there is intrinsic bias due to publication, "finding that estimates are related with their standard errors." However, OLS results can be influenced by studies that provide greater numbers of observations (equal weight is given to all observations). OLS estimates are representative of the appropriate sample frame since the selected observations are studies, not their resultant observations.[13] Stanley (2006) points out that paper's characteristics may reduce the biases when the model is re-estimated using proper meta-analysis weights. Model 3, a variation of the partial correlation coefficient affected by the estimation model specification still shows a positive and significant estimator. Interpreting these results can be misleading because we do not know the magnitude and direction of publication and other sources of bias.

Meta-analysis Figures B1 and B2 in the Annex 1 show the cumulative biases and weights from the sample. Both figures suggest that five papers contribute a larger proportion of the bias found in OLS estimations. The figures show how each paper affects meta-impact estimations. To correct for the bias the procedure estimates the weights needed to adjust regressions through a WLS in order to significantly correct such biases[14] (Costa-Font et al., 2011a). The effect size in Figure B1 shows that all papers have uniform distribution with exception of four papers that present standard error (S.E.) coefficients above 0.5, ranging from 0.55 to 1.40 with

[12] Models 1 and 2 estimate the determinants of a high t-value in impact estimates by including 2 distinct specifications using paper characteristics (e.g. precision variable). Based on Bruno and Campos (2011), we assessed whether the whole sample results might be the effect of a composition of very different types of papers, programs, and countries. Model 3 runs the determinants of paper correlations based on their characteristics. It measures the degree in which papers are differentiated enough—by sample choice, type of estimator, inclusion/exclusion of control, variables definitions –to have variability and a good representation of the literature (see Figures 1 and 2).

[13] In order to mitigate the influence of studies where the investigators report large numbers of observations the procedure must follow WLS estimations. Weighting is done by dividing the left-hand-side and right-hand-side variables of the regression by k_i, where k_i denotes the number of error observations from study i. The sum of the k_i for each study is one: the observations from studies that provide greater numbers of observations receive less weight in the estimation. Meta regression estimates calculate the weights adjusting publication bias as well (see Figure A2).

[14] The advantage of WLS in meta-analysis is that it assigns larger weights to those estimates with larger precision.

relatively large confidence intervals. These studies have large samples because they were part of large scale evaluations. This increases their influence in the total meta-analysis sample which might generate biases. However, these papers can also provide better precision in their estimates so that weights will compensate for both their relative sample size and their precision. Figure B2 shows a slightly different story. This figure shows the sources of effect size which come from publication bias.[15] The more papers are centered on unity, the more they are free of publication biases. For such papers, Meta-analysis procedures[16] allocate weights equal to zero. Other papers shifting slightly from unity and with large confidence intervals are imputed with very low weights. Papers with higher S.E. and relatively narrow confidence intervals, centered on unity, have the highest weights.

To corroborate the biases, Figures B3 and B4 in Annex 1 indicate which papers' characteristics are producing publication biases by type of evaluation method and data used. Not surprisingly biases relied on a paper that used a structural model and three papers that used qualitative methods. These same papers used ad hoc surveys and other complementary data to conduct the analysis. The methodologies that don't rely on robust estimates, and depend on calibrations or subjective criterion, produce publication biases by "using subjective or simulation methods where researchers adjust their models until the relationship between r and S.E. achieves some acceptable statistical significance" (Stanley, 2006).

Factors such as years of R&D program implementation, region and country of analysis, program targeting, and type of data contribute to isolate the program effect in terms of size and direction. Other paper-related characteristics are used to calibrate and adjust for such program effects and directions, given a paper's quality, methodology, citations, identification strategy, etc.

Table 4. OLS Estimates

OLS Estimates Impact Variables			
Dependent Variable	t-value of Estimator (Impacts)	t-value of Estimator (Impacts)	Study Correlation

[15] Even figures B3 and B4 in Annex 1 show the same number of papers inducing publication biases depending on the method or data used.
[16] These are computed using the metan, metabias and metacum stata commands in Stata.

Type data an Method			
Survey			-0.0281
s.e.			(0.03)
Panel			0.0179
s.e.			(0.03)
Cross-section			0.0521
s.e.			(0.04)
Region ECA		1.168	**0.0853****
s.e.		(0.95)	(0.03)
Region LAC		**1.131****	**0.0917****
s.e.		(0.45)	(0.04)
Region EU		2.746	**0.117****
s.e.		(1.71)	(0.04)
Region North America/Other		6.09	**0.107***
s.e.		(5.24)	(0.05)
Late Stage of Program			0.0281
s.e.			(0.03)
Quasiexperimental method used			-0.0436
s.e.			(0.03)
Complementarity			**0.0508****
s.e.			(0.02)
Paper precision	**0.118****	**0.109***	
s.e.	(0.06)	(0.06)	
Constant	**2.284*****		-0.0707
s.e.	(0.79)		(0.05)
Observations	*37*	*37*	*29*
R-squared	*0.158*	*0.45*	*0.429*

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

Once meta-graphs identify the papers that produce bias and assign weights to each paper, the WLS results show that the precision estimator is close to zero and the average R&D effect estimator from all paper estimates is 0.19. Results also show that by eliminating biases, the EU and North American papers showed highest precision, compared to the rest of the regions. Quasi-experimental studies with sample sizes above the average also showed the highest precision compared to other methods and sample sizes. In the area of R&D experimental evaluation designs are rare, so the second-best option for researchers is to rely on existing data to evaluate by statistically building a counterfactual.

Table 5 shows the truncated regressions (on the dependent variable) to verify the changes in coefficients with bias adjustment using WLS. OLS estimates produce positive but insignificant coefficients for the firm-level data dummy variable and for the dummy indicating paper focusing on innovation impacts measurement.

These are two important variables because the former refers to estimating impacts with the proper measurement unit (firm) and the latter conveys innovation activities where firms invest subsidized resources. The statistically insignificant coefficient in both variables would suggest that both factors do not contribute to enhance heterogeneity between papers and thus these factors become irrelevant for the meta-analysis perspective. However, once the weights are estimated for each paper, minimizing sources of biases given their characteristics, the coefficient of both variables becomes statistically significant at least 95 percent. The negative sign of the variables that also turn statistically significant (years of program implementation, program investment amount, and peer review publication) suggest a high degree of heterogeneity, in spite of the truncation of the model.

Table 5. Truncated Regression (Minimum Correlation)

	Study Correlation (1)	**Study Correlation** (2)
Firm Level Data	0.0356	**0.299*****
s.e.	(0.07)	(0.10)
Identification 2 stage	-0.053	**-0.0486***
s.e.	(0.05)	(0.03)
Identification Dif in Dif	-0.0273	-0.0176
s.e.	(0.09)	(0.05)
Identification PSM	-0.0873	**-0.0547****
s.e.	(0.05)	(0.03)
Identification RD	-0.105	-0.0348
s.e.	(0.11)	(0.06)
Identification FE/RE	0.0611	0.0198
s.e.	(0.05)	(0.02)
Identification IV	-0.0144	-0.00475
s.e.	(0.06)	(0.03)
Dummy small business support	-0.00493	-0.0244
s.e.	(0.05)	(0.03)
Dummy SME	0.0327	0.0155
s.e.	(0.04)	(0.02)
Log years program	-0.0567	**-0.0416****

implementation		
s.e.	(0.04)	(0.02)
Log investment (USD,PPP)	-0.0221	**-0.0236****
s.e.	(0.02)	(0.01)
Log number citations	0.00882	0.00532
s.e.	(0.02)	(0.01)
Region Categorical Variable	0.00171	0.00155
s.e.	(0.01)	(0.00)
Log Rate Quality paper	0.119	0.0786
s.e.	(0.17)	(0.09)
Dummy peer reviewed pub.	0.0138	**-0.105*****
s.e.	(0.07)	(0.03)
Dummy Paper focus innovation	0.0593	**0.0343****
s.e.	(0.05)	(0.02)
Dummy outcome indicators incl.	-0.0146	-0.0193
s.e.	(0.05)	(0.03)
Dummy Robustness Checks	-0.026	0.00312
s.e.	(0.05)	(0.03)
Constant	-0.0743	-0.215
	(0.30)	(0.16)
Observations	*28*	*28*
Pseudo R2	*0.661*	*0.656*

(1) Simple truncated model low correlation
(2) Simple truncated model with precision weights
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Source: Own estimations

Harbord ‘s test (Table 6) confirms the presence of small study effects bias but these are not higher than the actual effects adjusted by reciprocal sample sizes.[17] The Harbord test p-value rejects the null hypothesis, indicating the presence of small study effects and bias. This requires further correction by estimating the WLS with meta-regression adjustments (Table 7). The variable that measures the additionality effects is positive (0.195) and significant at the 99

[17] Egger et al. (1997) proposed a test for asymmetry of the funnel plot. This is a test for the Y intercept = 0 from a linear regression of normalized effect estimate (where estimate is divided by its standard error) against precision (reciprocal of the standard error of the estimate). Harbord et al. (2006) developed a test that maintains the power of the Egger test whilst reducing the false positive rate, which is a problem with the Egger test when there are large treatment effects, few events per trial, or when all trials are of similar sizes. The original Egger test should be used instead of the Harbord method if there is a large imbalance between the sizes of treatment and control groups. However, the Harbord test has an advantage because it regresses $Z/\sqrt{V}$ against $\sqrt{V}$, where Z is the efficient score and V is Fisher’s information (the variance of Z under the null hypothesis).

percent level, effectively indicating that the additionality of R&D is present in the whole set of studies. The negative sign of the Standard Error of the precision indicates that studies with higher SE are predominant, although this coefficient was not statistically significant.

Finally, the WLS meta-analysis revealed that R&D evaluations show that public funds do not crowd out but incentivize firms to revert funds into R&D. However, the small sample size of surveyed studies produces high ranges of confidence intervals across independent variables coefficients across different models. The high heterogeneity of precision is explained by the wide variety of methodologies used to estimate impacts. The use of "gold standard" evaluation methods (randomized assignment) is not common at all in the literature sample.

Table 6. Meta Regression Precision, Study Characteristics and Study Effects tests

a. Precision

Meta-regression				Number of obs	=		37
REML estimate of between-study variance				tau2	=		51.89
% residual variation due to heterogeneity				I-squared_res	=		100.00%
Proportion of between-study variance explained				Adj R-squared	=		10.71%
With Knapp-Hartung modification							
Single test for Precision (against paper correlation)	**exp(b)**	**Std. Err.**	**t**	**P>|t|**		**[95% Conf. Interval]**	
Precision	1.12	0.06	2.2	0.034		1.009	1.242
Test for residual between-study variance (of tau2=0) Q_res (35df)					=		2.50E+06
				Prob > Q_res	=		0
Likelihood-ratio test of tau2=0: chibar2(01)				Prob > chibar2	=		0

Source: Own estimations

b. Study Characteristics

Meta-regression				**Number of obs**	=	**29**
REML estimate of between-study variance				tau2	=	0.001511
% residual variation due to heterogeneity				I-squared_res	=	8.25%
Proportion of between-study variance explained				Adj R-squared	=	59.16%
Joint test for all covariates				Model F(10,18)	=	2.06
With Knapp-Hartung modification				Prob > F	=	0.0879
pcorr	**exp(b)**	**Std. Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
Survey Used	0.96	0.03	-1.31	0.21	0.89	1.03
Panel	1.02	0.04	0.41	0.69	0.94	1.10
Cross-section	1.01	0.10	0.07	0.94	0.82	1.24
Region ECA	1.08	0.07	1.14	0.27	0.94	1.24
Region LAC	1.12	0.14	0.89	0.39	0.86	1.46
Region EU	1.18	0.08	2.47	0.02	1.03	1.36
Region North America/other	1.22	0.09	2.76	0.01	1.05	1.43
Late Stage	1.01	0.04	0.26	0.80	0.93	1.10
Quasiexperimental method used	0.93	0.04	-1.91	0.07	0.86	1.01
Complementarity	1.06	0.04	1.68	0.11	0.99	1.14

Source: Own estimations

c. Study Effects

Harbord's modified test for small-study effects:					
Regress Z/sqrt(V) on sqrt(V) where Z is efficient score and V is score variance					
Number of studies = 31	Root MSE = .5442				
Z/sqrt(V)	**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf. Interval]**
sqrt(V)	1.04	0.32	3.25	0.00	0.38
bias	0.89	0.14	6.50	0.00	0.61
Test of H0: no small-study effects	*P = 0.000*				

Source: Own estimations

Table 7 Meta-Regression with Weights for Bias Correction

Study Correlation Coefficient	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
precision	**0.001**	**0.001**	**1.410**	**0.189**	**0.000**	**0.002**
Estimator	**0.195**	**0.040**	**4.840**	**0.001**	**0.105**	**0.285**
Standard Error of Precision	**-0.057**	**0.214**	**-0.270**	**0.794**	**-0.533**	**0.419**
Log Total Investment R&D Prog.	0.002	0.005	0.390	0.704	-0.008	0.012
Log Number of web views	-0.003	0.005	-0.550	0.596	-0.014	0.009
Log Study Sample Size	-0.008	0.009	-0.860	0.412	-0.028	0.013
Log Years Implementation	-0.018	0.016	-1.130	0.283	-0.055	0.018
Survey Used	-0.003	0.012	-0.210	0.834	-0.029	0.024
Panel	-0.031	0.020	-1.500	0.164	-0.076	0.015
Cross-section	0.023	0.050	0.470	0.652	-0.088	0.135
Region ECA	0.028	0.025	1.130	0.285	-0.028	0.085
Region LAC	0.059	0.034	1.730	0.115	-0.017	0.135
Region EU	0.028	0.026	1.070	0.309	-0.031	0.087
Region North America/other	0.081	0.047	1.740	0.112	-0.023	0.185
Late Stage	-0.008	0.021	-0.370	0.721	-0.054	0.039
Quasiexperimental method used	-0.029	0.017	-1.730	0.115	-0.067	0.009
Complementarity	0.034	0.019	1.790	0.094	-0.009	0.076
Constant	0.054	0.109	0.490	0.634	-0.190	0.298

Source: Own estimations

One thing that is worth noticing is that this 0.195 average effect of additionality is estimated with Meta-Regression correcting with WLS for papers that induce bias. The literature suggests estimating the effects with fixed and random effects. Assuming that the n studies provided have an (statistically) accurate average additionality R&D effects, and because each study provides its own S.E. of this estimation coefficient, then we need to take assumptions on the effect changes according to size of study or an attribute (e.g. region, country). However, fixed effects will assume a single size effect for each paper. When working with papers that have different population targets and also different sample sizes used to estimate R&D additionality impacts, then random effects is suited best since it allows capturing different sizes, although we cannot separate the sample size and population target effects. An important element about our results is that observed paper characteristics induce a higher change in the meta-coefficient than when using fixed or random effects, leading to conclude about the presence of relatively smaller size effects and biases.

The different specifications shown in Table 8 demonstrate that weighting observations from meta-analysis substantially reduces heterogeneity. This implies that the coefficients are relatively similar between different types of specifications and methods used, keeping the rest of paper and evaluation methods characteristics constant. The coefficient of additionality impacts on R&D ranges from 0.166 to 0.252, with reasonable confidence intervals at the 95 percent level. The number of R&D impact evaluations, which predominantly target North American and European countries, are limited and such a small number constrains the robustness of the results. As more evaluations are conducted with more robust techniques, more can be said about the direction of the effects with statistical confidence. At this point, the analysis revealed that based on the surveyed studies there is indication of positive R&D impacts, but this effect should be reconfirmed with more rigorous impact evaluations. The estimation results have a large range, suggesting that the meta-data are highly sensitive to the method used. Clearly, the meta-analysis methods corrected biases into a certain extent, but the data is not as rich as other meta-analysis studies that rely on hundreds of studies.

Table 8

Meta-Analysis of Additionality of Subsidies on R&D Investments

Using Weighted Least Squares correcting for Publication Bias

	Estimate	Lower Bound (95%)		Upper Bound (95%)
WLS	0.252	0.193		0.312
WLS Full Specification	0.195	0.105		0.285
Fixed Effect (Region)	0.194	0.101		0.271
Random Effects	0.166	0.140		0.192

Source: Own estimations

4. Concluding Remarks

R&D evaluation studies are not characterized by having strong and rigorous methodologies. However, data from surveyed studies are heterogeneous enough to allow for meta-data analysis. Very few studies are devoted to quantitatively estimating the impact of innovation, although there is a growing body of literature on the effects of direct R&D grants. Despite this, there is still a knowledge gap on how input, output and outcomes relate to each other in the presence of grants and subsidies (Veryzer, 1998). Some studies that claim to be evaluations fail to fill the gap of endogeneity between R&D interventions and a firm's outcome variables, in impact studies. Without solving for endogeneity, impact results are biased and in some cases spurious.

The literature review showed that results vary substantially depending on the countries and/or the types of sectors/industries analyzed; other categories of R&D interventions are inconclusive due to the nature of the weak designs created to evaluate these interventions. Overall, the meta-analysis revealed that there are some studies that produce biases in meta-estimates: such studies tend to use perception surveys and subjective methods to evaluate R&D programs. The meta-analysis bias correction assigns weights to be computed in meta-regressions (WLS), where the papers producing bias receive a weight of zero. The final regressions with the bias-correction weights show positive and significant impacts of R&D subsidies on a firm's innovation activities with a mean of 19 percent (compared to different types of non-recipient/counterfactual firms) controlling for precision, methodology, and paper attributes. Although the results are robust from a meta-analysis standpoint, the weakness of the original methodologies makes it hard to build a case for causality and calls for increasing randomized designs in R&D interventions.

The design of a public program of subsidies to business R&D projects requires defining a selection and ranking system in order to decide which projects should be supported. The decision criteria of the agency should be part of the evaluation of a public program because, as the structural models on this subject show (David et al., 2000), such decision criteria in the selection of projects has as major impact on the results of the program as well as on the effect of the subsidies.

Government-wide evaluations of efficiency are often based on complex composite indicators. These indicators are useful to get a broad overview of efficiency gains achieved. However, in order to arrive at concrete policy recommendations, it is more promising to investigate the efficiency of public expenditure in individual spending areas. Growth enhancing expenditures, such as R&D, education and to some extent infrastructure, as well as expenditures, affected by the ageing of population (such as health care), are first candidates for such investigations.

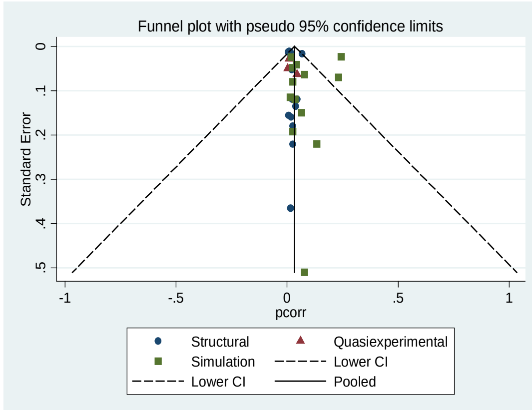
Overall, the quality of the evaluations is also affected by the fact that selection of the funded projects is not often granted based on quality and on specific goals. Despite that it is recognized that rigorously evaluating R&D interventions has high implementation and monitoring costs- there is a new set of evaluations that are introducing more robust statistical analysis to build valid counterfactuals. Most of the time, randomized methodologies are not to be applied to these sorts of interventions because of competition rules and the institutional regulations (Lentile and Mairesse, 2009). But more often, governments sponsor cutting-edge pilot programs that can be subject to prospectively-designed randomized evaluations. With these types of methods the researchers may even have more outcomes to evaluate, by building different comparison groups. More recently, researchers are intrigued by determining if a lower user cost entails higher R&D expenses, or the degree to which innovation outputs and productivity relate to one another in the presence of subsidies. Other important research questions deal with understanding at which point R&D produces desired effects. If marginal productivity of R&D is decreasing, additional units could generate less innovation. Further, it is important to include R&D evaluation programs in the innovation agenda by knowing the degree to which subsidies stimulate innovation that is valued by the market.

As the evaluation methods used are more sophisticated and statistically robust, the effects appear to be conclusive. Recent R&D evaluations that follow a logical framework and focus on addressing the issue of causality, through matching methods, find consistently positive and significant results. A tendency to overestimate impacts in the absence of statistically valid counterfactual, regardless the aspect of the firm influenced by R&D subsidies reflects the ambiguous results come from a design effect of the evaluation methods used. The meta-analysis revealed that there is indication of positive R&D impacts, but such effects are highly sensitive to the method used. In addition, evaluations of certain areas of R&D still have important knowledge gaps because of lack of data.

Finally, one of the main challenges to overcome the insufficient number of randomized evaluations has to do with the potential endogeneity of the subsidy, the assignment of which fails to satisfy the randomness property that should characterize pure social experiments. This is why most recent R&D IEs studies estimate a counterfactual through Propensity Score Matching. An evaluation of the expected innovative outcome, by both the firm, which has to decide whether to apply for the subsidy, and the public agency, which must decide which projects to subsidize, is likely to precede the allocation process. This makes public funding an endogenous variable with respect to innovation itself. This is one of the reasons why evaluators tend to (over)estimate impacts, using existing (administrative) data. But these types of models fail to correct for important source bias. Prospectively designed impact evaluations help researchers incorporate other factors that produce biased estimates. This is important because the scale of intervention, size of the firm, type of beneficiary, and project attributes can produce biases in R&D impact estimates.

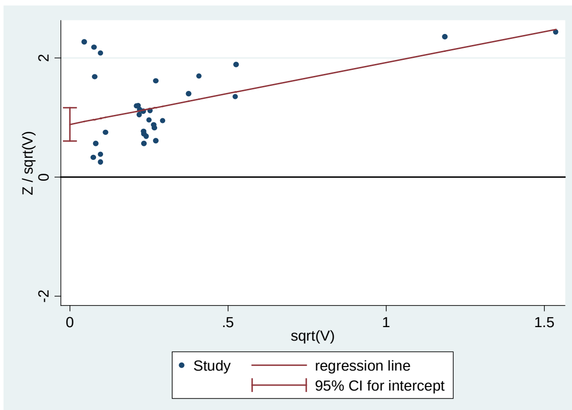
ANNEX 1 – Additional Figures

Figure A1. Most IE Methods used are Comparable



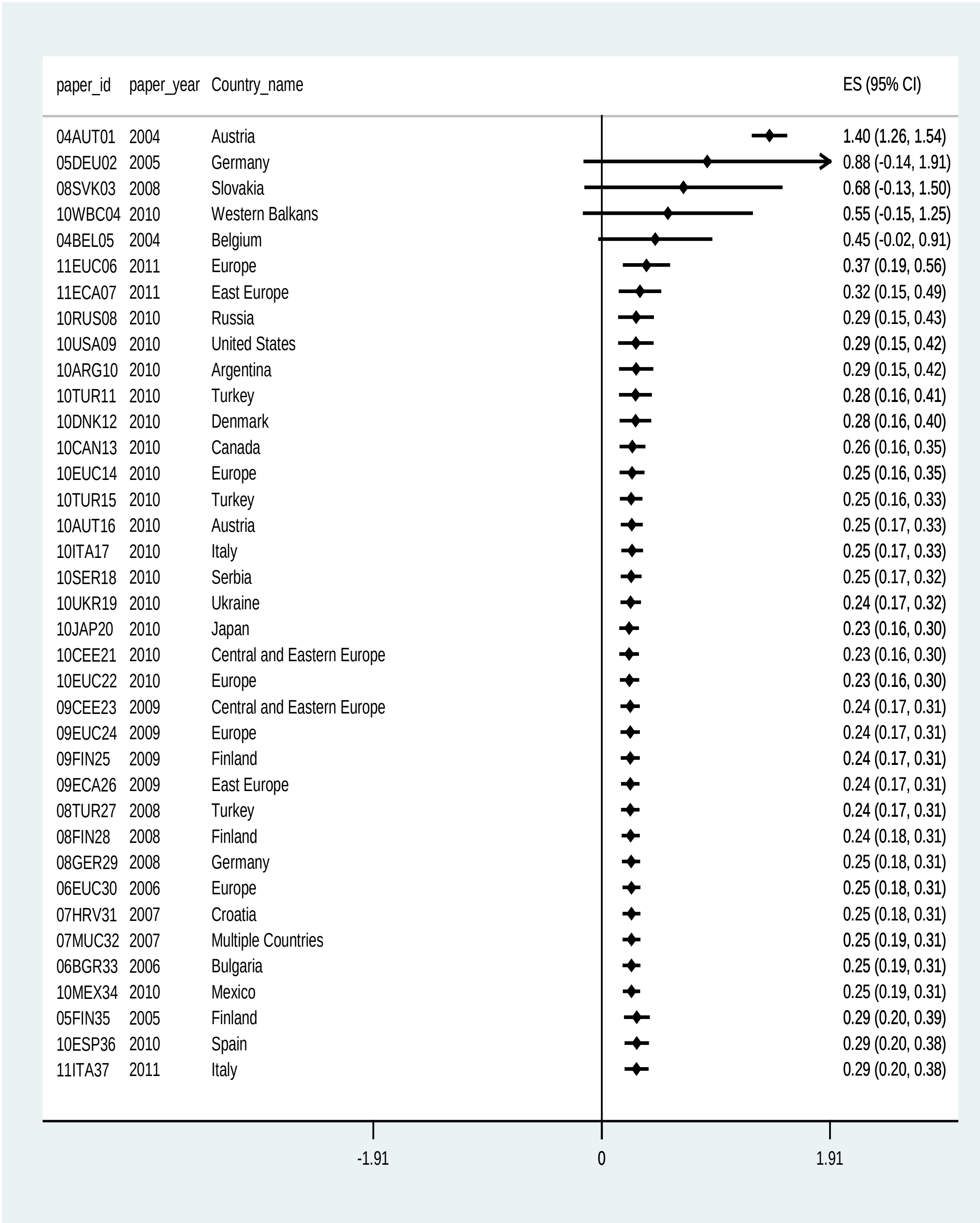
Source: Own estimations

Figure A2. Meta-regression Shows Some Publication Bias



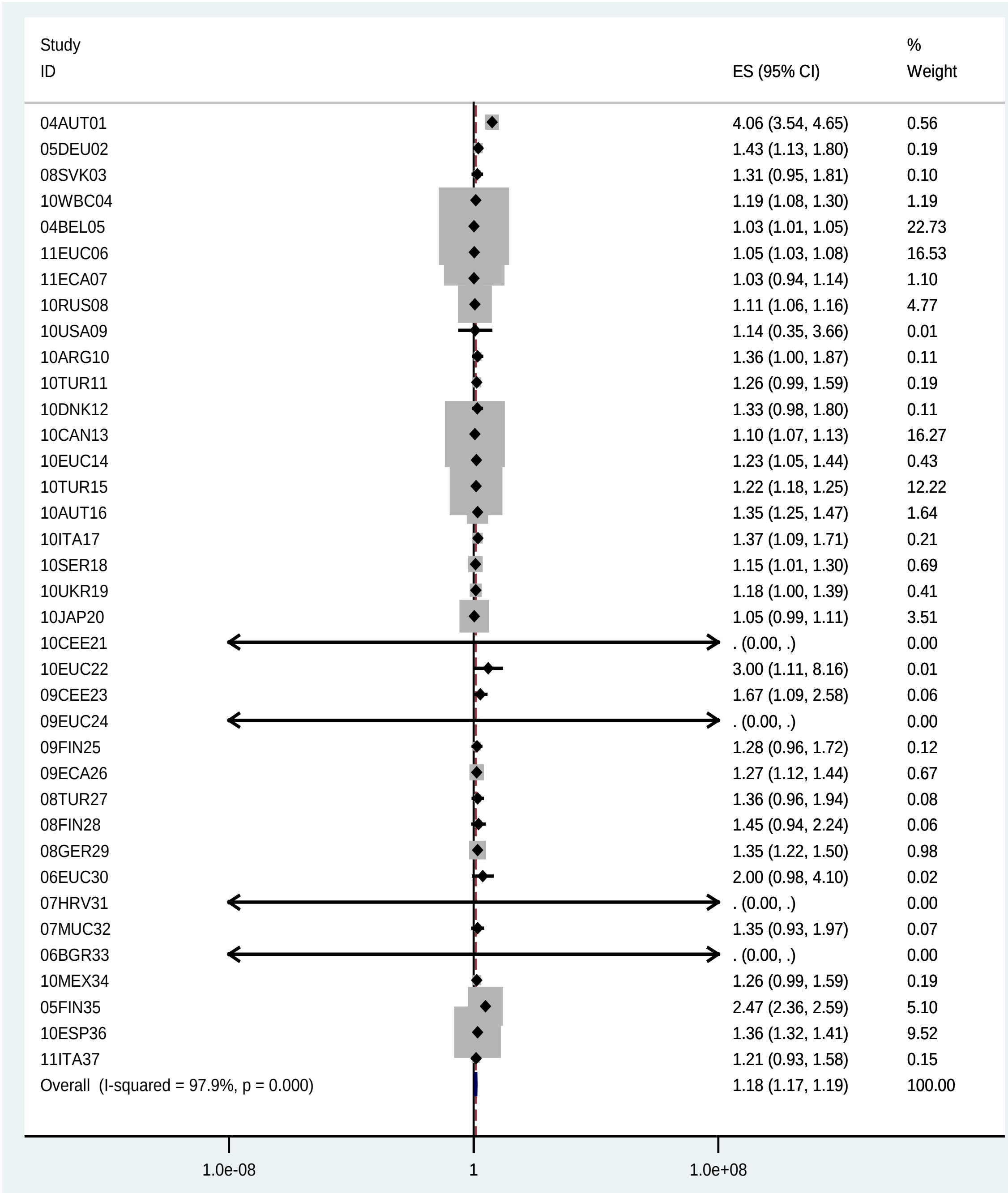
Source: Own estimations

Figure B1. Sources of Cumulative Publication Bias



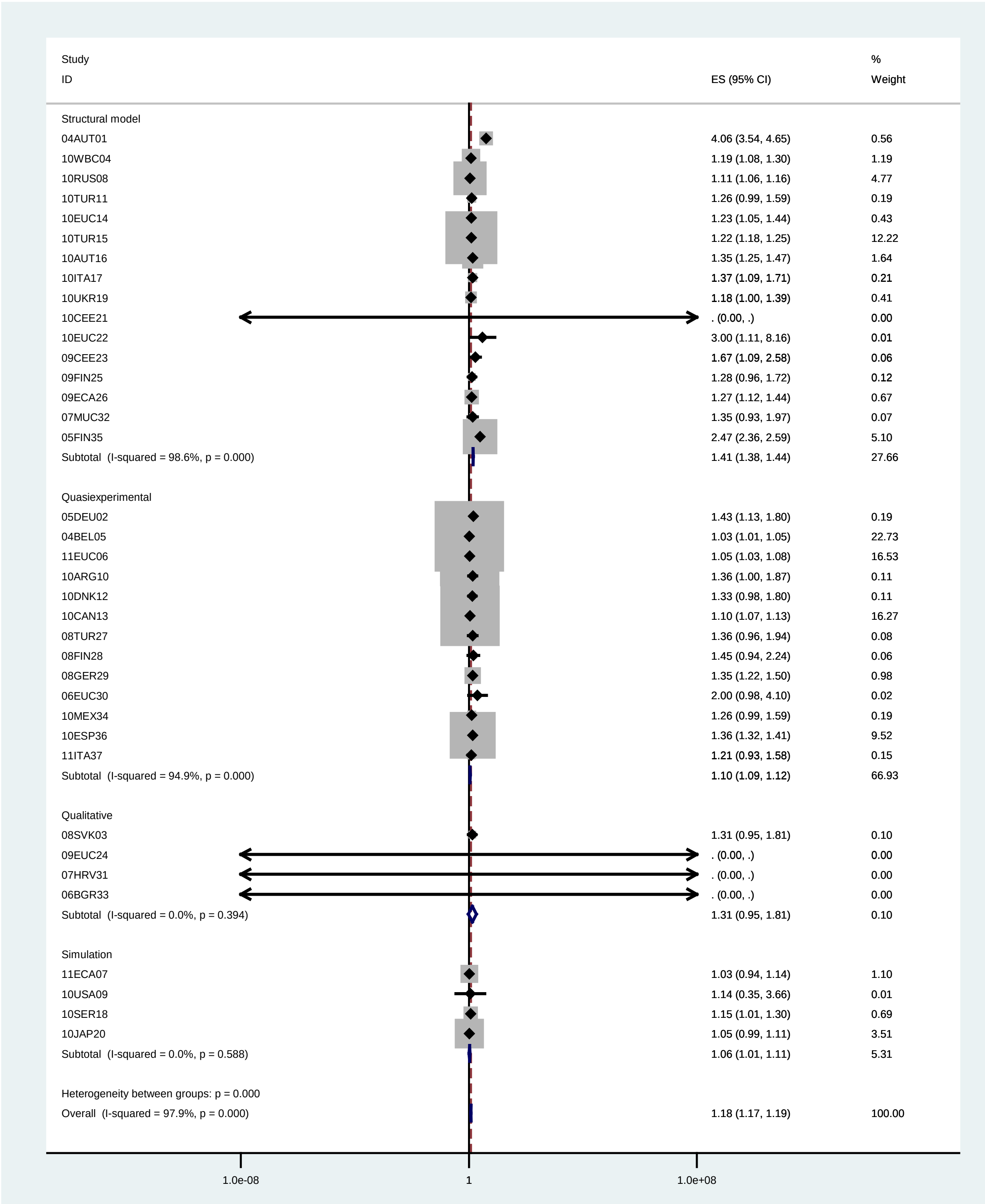
Source: Own estimations

Figure B2. Papers that Produce Publication Bias Affect Weights for Meta Impact Estimations



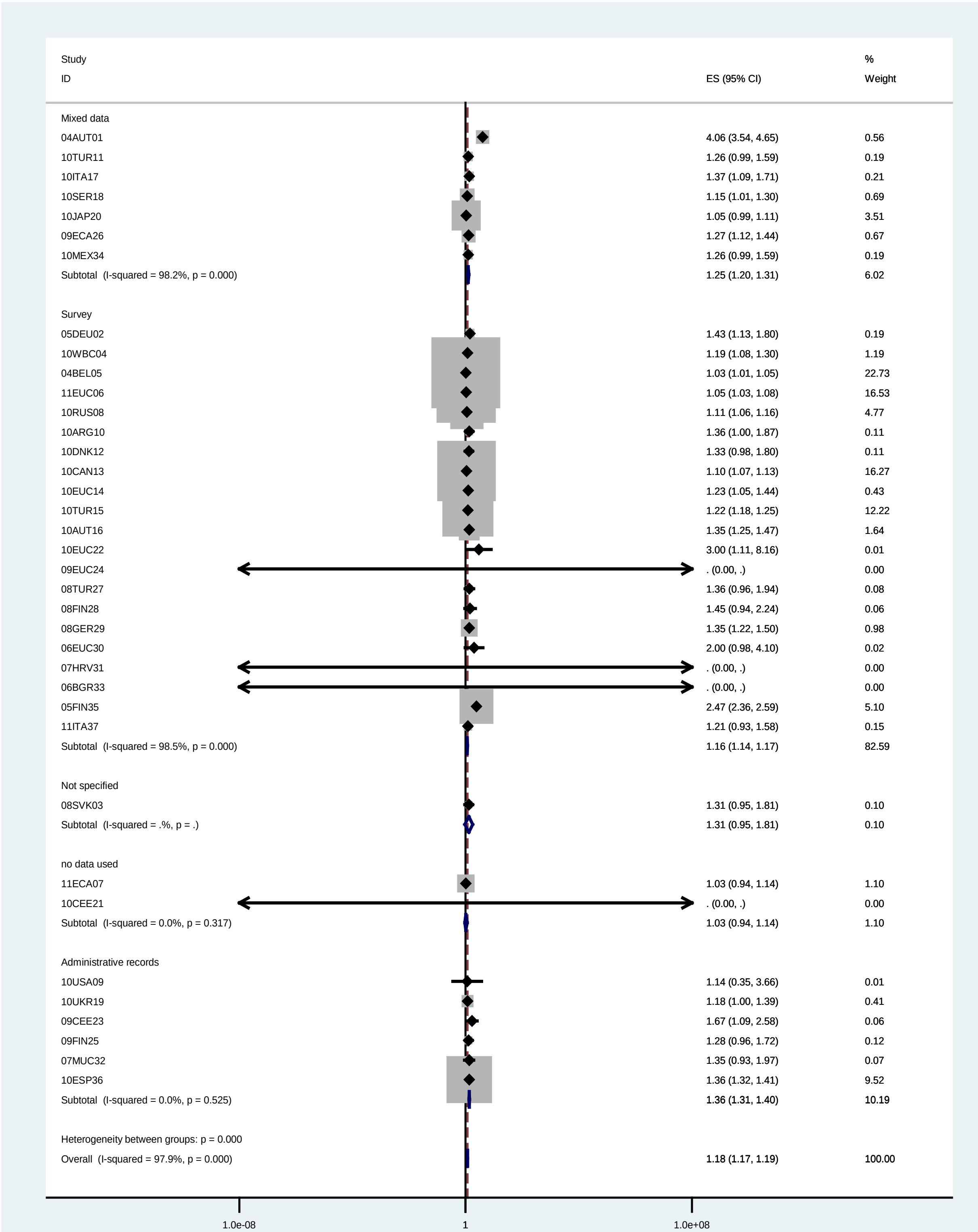
Source: Own estimations

Figure B3. Papers with Qualitative, Structural or Other Non-specified Methods Produced Biases



Source: Own estimations

Figure B4. Papers with Publication Biases used Ad Hoc Surveys to Evaluate R&D Additionality



Source: Own estimations

Annex 2: Methodological Appendix

Structural Modeling for R&D Evaluations

The existing treatment evaluation literature offers alternative methodologies to deal with such potential endogeneity, however each impose restrictive conditions. In particular, these approaches rely on the hypothesis that depending on a set of observable explanatory factors X, the alternative outcomes y(1) (with treatment) and y(0)(without treatment) are orthogonal to the treatment (D):

$$y_0, y_1 \perp D|X$$

These approaches neglect the possibility that observable factors may simultaneously affect both the treatment (D) and the adopted performance measure (y). Simultaneous equation systems accomplish this aim, jointly taking into account the treatment assignment process and its outcome, i.e. checking whether the funding allocation process is partially determined by the same factors affecting the innovative process (endogeneity). In this framework, an endogenous dummy variable (D) becomes the dependent variable of a participation equation where the subsidy can be explained by the same factors affecting a firm's innovative performance (Busom, 2000). In other words, two different regimes for the innovative performance are allowed, public support playing the role of endogenously switching firms from one regime to the other. Therefore, the resulting switching model can be written as:

$$\begin{cases} D_i^* = \alpha' z_i + \mu_i; D_i = 1 \; if \; D_i^* > 0, 0 \; otherwise \\ y_{1i} = \beta_1' x_i + \varepsilon_{1i} \qquad\qquad \varepsilon_{1i} \sim N(0, \sigma_{11}) \\ y_{0i} = \beta_0' x_i + \varepsilon_{0i} \qquad\qquad \varepsilon_{0i} \sim N(0, \sigma_{00}) \end{cases}$$

$$corr[\mu_i, \varepsilon_{1i}] = \rho_{\mu 1}; corr[\mu_i, \varepsilon_{0i}] = \rho_{\mu 0}$$

where the set z of factors determining D partially overlaps the set x that explains the innovative outcome level y; the last row accounts for the likely correlation between the treatment-equation and the performance-equation error terms (endogeneity). Such a simultaneous model fulfills two needs: firstly, it allows us to correct for funding endogeneity, producing consistent estimates of the performance equation (separately estimated on the two sub-samples of treated and non-treated firms), and secondly, it solves the missing-data problem affecting the treatment evaluation literature. Indeed, although we cannot directly observe how supported firms would have behaved had they not received the subsidy, we can nevertheless estimate the relevant model on the non-supported firms. The average treatment effect on treated firms can thus be computed consistently as:

$$ATET = E[y_{1t}|x, D_t = 1] - E[y_{0t}|x, D_t = 1]$$

where the estimated coefficients obtained using the sub-sample of non-supported firms are applied to the supported ones, in order to achieve an estimate of the potential productivity the supported firms would have reached had they not received the subsidy. This approach here is further developed in order to take into account a second source of endogeneity arising from the possible simultaneity between government intervention and the qualitative composition of the innovative output. Indeed, while receiving a subsidy is likely to foster one innovative typology at the expense of others, it appears equally plausible that the qualitative composition of the innovation a firm has realized may affect the probability of receiving such a subsidy. This two-way simultaneous relationship should be taken into account when correcting for the selection of product innovators only. This is why we replace the participation equation identifying the switching in the standard endogenous switching models, with a bivariate model. Therefore the estimated bivariate switching model will be:

$$\begin{cases} funding_i^* = \alpha_a' z_{ai} + \mu_{ai}; funding = 1 \; if \; funding^* > 0, 0 \; otherwise; \\ PDT\_ONLY_i^* = \alpha_b' z_{bi} + \mu_{bi}; PDT\_ONLY_i \;\; = 1 \; if \; PDT\_ONLY_i^* > 0, 0 \; otherwise \end{cases}$$

$$PDTV_i \begin{cases} \beta_{11}' x_i + \varepsilon_i \; if \; funding = 1 \; and \; PDT\_ONLY \;\; = 1 \\ \beta_{01}' x_i + \varepsilon_i \; if \; funding = 0 \; and \; PDT\_ONLY \;\; = 1 \\ \beta_{10}' x_i + \varepsilon_i \; if \; funding = 1 \; and \; PDT\_ONLY \;\; = 0 \\ \beta_{00}' x_i + \varepsilon_i \; if \; funding = 0 \; and \; PDT\_ONLY \;\; = 0 \end{cases}$$

The first system thus accounts for the "double switching" (i.e. the joint probability of getting the subsidy and of engaging in product innovation only) that endogenously affects the productivity equation (second system). ε, ua and ub follow a trivariate normal distribution with variances σ2, 1 and 1 respectively, and correlations ρab, ρεa and ρεb defined as follows:

$$\rho_{ab} = corr(u_a, u_b); \rho_{\varepsilon a} = corr(u_a, \varepsilon) = corr(u_b, \varepsilon)$$

The first two selection equations can thus be correlated with each other besides each being individually correlated to the main productivity equation. This fully incorporates the correction for the product-only sample selection into the bivariate switching model. Of course, once a bivariate (rather than an univariate) selection is implemented, four instead of just two different regimes are identified, accounting for the potential specificities that characterize each possible combination of the two switching variables: (1, 1); (0, 1); (1, 0) and (0, 0).

From a computational point of view, four productivity equations should be estimated, each of them augmented by two additional terms (inverse Mills ratios) correcting for the double selection bias. Thus, for instance, focusing on the sub-sample identified by the combination (funding=1 & PDT_ONLY=1), the estimated performance equation will be:

$$PDTV_i = \beta_{11}' x_i + \theta_a \lambda_a + \theta_b \lambda_b + \varepsilon_i$$

where:

$$\theta_a = \sigma \rho_{\varepsilon a}; \theta_b = \sigma \rho_{\varepsilon b}$$

$$\lambda_a = \phi(\omega_a)\,\Phi\,[\omega_b - \rho_{ab} funding)/(1-\rho_{ab}^2)^{\frac{1}{2}}]/\,\Phi_2$$

$$\lambda_b = \phi(\omega_b)\,\Phi\,[\omega_a - \rho_{ab} PDT\_ONLY)/(1-\rho_{ab}^2)^{\frac{1}{2}}]/\,\Phi_2$$

Where the lambda equations are obtained through probit estimations. The same procedure applies to the other three sub-samples. For our purposes, the relevant ATET will be:

$$E\left[PDTV_{1i}\,\middle|\,x, funding_i = 1\&PDT_{ONLY} = 1\right] \cdot\ E\left[PDTV_{0i}\,\middle|\,x, funding_i = 1\&PDT_{ONLY} = 1\right]$$

where, following the same procedure is adopted for the univariate endogenous switching model. The coefficients obtained on the sub-sample of non-supported product innovators will be applied to the supported ones in order to obtain an estimate of their potential productivity had they not received the subsidy (counterfactual).

Meta-Analysis Effect Sizes

The quantitative relationships of interest may be univariate, bivariate, or multivariate. Nearly all meta-analysis is done on bivariate relationships, e.g., treatment-control differences on outcome variables or covariation between two variables. There are three main families of effect size statistics for bivariate relationships:

1. Relationships between two continuous variables, e.g., score on a personnel selection test predicting a job performance measure; risk or diagnostic measure predicting later outcome; concurrent relationship between SES and political attitudes.

Product-moment correlation:

$$r_{xy} = \frac{\sigma_{xy}^2}{\sigma_x \sigma_y}$$

Computational form (Fisher's Z transform):

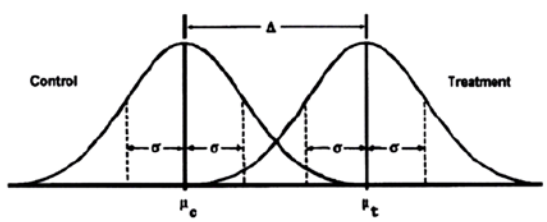
$$r_Z = .5\log_e\left[\frac{1+r}{1-r}\right] \qquad \text{Single study fixed effect standard error} \quad SE_r = \frac{1}{\sqrt{N-3}}$$

2. Relationships between one dichotomous and one continuous variable; e.g., difference in R&D and firms' outcomes between comparison groups.

$$d = \frac{\overline{X}_1 - \overline{X}_2}{\sigma_{pooled}}$$

Computational form (small sample bias correction; Hedge's *g*)

$$g = \left(1 - \frac{3}{4N-9}\right) \times d \qquad \text{Single study fixed effect standard error} \quad SE_g = \sqrt{\frac{n_1+n_2}{n_1 n_2} + \frac{g^2}{2(n_1+n_2)}}$$



Effect Size Estimation and Adjustments to the Estimates

Effect sizes from different studies are based on different sample sizes; those based on large samples should be given more weight in the analysis. All analysis with effect sizes is weighted analysis with the effect sizes weighted by an index of statistical precision (sampling error).

Fixed effects: Effect sizes from the studies are assumed to all are assumed to estimate the same population effect size, so there is no study-level sampling error. Or, alternatively, the whole effect size population of interest is represented so that no sampling was done; i.e., no inferences will be made to any studies/effect sizes absent from the set under consideration.

The standard error for an individual effect size is the respective SE value from the formulations shown above. In any analysis, each effect size is weighted by the inverse sampling error variance 1/SE.^{2} Inferential statistics for those analyses are based on the assumption of no between-study sampling variance.

ANNEX 3. List of Papers/Publications Used in Meta-Analysis		
Authors (Year)	**Intervention/objective**	**Methodology/Design**
Lopes Bento (2011)	Public funding targeted to internal R&D investment and to total innovation intensity	Matching estimators using cross-country and firm level. Sample of 9790 observations of 5 different countries, out of which 3854 received R&D subsidies. Non-experimental.
Varga and Veldt (2011)	Fields of intervention of R&D schemes: Infrastructure, Agriculture, RTD, HR, TA	The model belongs to the class of micro-founded dynamic general equilibrium (DGE) models used in economic policy institutions. The model employs the Dixit-Stiglitz product-variety framework and the mechanism through which this R&D spending supports growth in the model: by reducing costs, the cohesion programme spending makes it easier for new start-ups to enter the market and so support the introduction of new products. Non-experimental.
Akhmedjonov (2010)	Address the R&D policies and its importance to various measures of human capital, financing and competition environment in the process of technology diffusion	Uses a symmetric Cobb-Douglas preferences model with credit-constrained firm with resources (assets and available credit). hypothesis of this research is that human capital development is complementary to innovation and technological change. Panel data estimation. Non-experimental.
Akcigit and Kerr (2010)	R&D exploration and exploitation innovations interventions impact on economic growth	Model that incorporates the empirical regularity that exploration R&D does not scale as fast as exploitation R&D with firm size. Study the implications of program heterogeneity on the R&D, innovation, and growth dynamics of firms. Non-experimental.
Wallace (2010)	European public research and development (R&D) subsidies that support precompetitive development	EU framework programs. Non-experimental.

	(PCD)	
Lopez, Reynoso y Rossi (2010)	FONTAR Program funds projects presented by private firms which aim at improving their competitive performance through technological innovation activities	Test the Additionality versus Crowding Out hypothesis. That is, we will evaluate whether the presence of the public aid to innovation complements or crowds out subsidized firms' investments in innovation activities. Matching estimators used to evaluate. Non-experimental.
Erden (2010)	Academic research projects that are supported by TÜBİTAK under Academic Research Funding Programs Directorate and under Basic Sciences Research Funding Group. Funds administered by Grant Committee.	Assesses social benefit of physics projects supported by TÜBİTAK. Cost-benefit analysis, based on IE. Non-experimental.
Marino, Parrota and Sala (2010)	Danish R&D grant support system, assess the "additionality" effects.	Categorical and continuous treatment schemes are used as alternatives to the traditional binary approach. Non-experimental.
Baghana (2010)	Compare and Assess the input additionality of R&D subsidies in Quebec. Explore how firms of different technological level respond to public grants	Model: Simple Cobb-Douglas production function with labor productivity function embedded. Identification strategy: Conditional semiparametric difference-in-differences estimator (CDiD) allows not only selection on both observables and on unobservables, but it also resolves the multidimensional and contemporaneous heterogeneity problems. Non-experimental.
Roper (2010)	European Charter for Small Enterprises in 2003 the Western Balkans Countries. ECSE has in place the basic legal and regulatory frameworks necessary for entrepreneurship and business development	Model of innovation or knowledge production function (Griliches 1992; Love and Roper 1999). Bivariate probit models of the innovation production function, reflecting the probability that locally-owned firms undertook either (new or upgrading) product or service innovation during the 2002 to 2005 period Non-experimental.

Bayona-Sáez and Garcia-Marco (2010)	R&D Eureka grants achieve competitive gains as a first step towards higher profitability. Participation in a Eureka Program research project will have a positive effect on performance in participating firms.	GMM-Arellano Bond Estimators, with random subsample for control group.
Bascavusoglu-Moreau and Colakoglu (2010)	Investments in innovation promoted through tax incentives, matching grantsand reimbursable loan schemes from National Systems of Innovation. Analysis of evaluating the determinants of Turkish SMEs' innovative capabilities	Knowledge production function (Griliches, 1979), whichmodels the "functional relationship between the inputs of the knowledge production and its output that is economically useful new technological knowledge". The patents do not play any explicit economic role in Griliches' model. They are just an indicator of innovative activity. OLS model used corrected for heteroskedasticity and clustering FE. Non-experimental.
Garcia and Mohnen (2010)	Austria incentives are only granted for eligible expenditures, i.e. those that are considered as valuable to the economy	Structural model explaining the determinants of various sources of government support and their effects on R&D and innovation output. Government support, R&D and innovative sales are all three endogenous. Estimation made through asymptotic least squares using 2 stages (probit and tobit). Non-experimental.
Cerulli and Poti (2010)	Fondo per le Agevolazioni della Ricerca (FAR) managed by the Italian Ministry of Research that is one of the two main pillars over which national R&D and innovation supporting policies are based. Contains both bottom-up and top-down measures as well as basic and more applied research projects.	Measuring the presence or absence of additionality. Structural model identifies the optimal level of R&D investment as the point in which marginal rate of returns (MRR) and marginal capital costs (MCC) associated to R&D investments. Econometric methodology used to evaluate the input (R&D outlay) and output (patents) additionality of the FAR fund is based on the literature on "program evaluation". Models estimated with OLS pooled sample. For the output additionality equation a Poisson regression was used. Non-experimental
Simon (2010)	Assess R&D policies at the macro and manufacturing levels	Simple growth model, using pooled OLS. Non-experimental.

Almus and Czarnitzki (2003)	Eastern German firms which receive public R&D funds.	Non-parametric matching to identified Counterfactual with unbounded propensity score (distribution approximation). Random replacement samples drawn to check robustness. Non-experimental
Brown et al. (2010)	Test how ownership type affects the propensity to invest in R&D in Ukraine	First step Tobit regressions to explore how ownership and firm origins are related to the intensity of different investment types. Second set of regressions calculates the productivity growth returns to different investment types (Neoclassical production function). Non-experimental.
Matsumoto et al. (2010)	Public research Institutes R&D grants. Study economic impacts for industry and sectors.	Ad hoc design based on modeling and simulation. Case studies for qualitative assessment. Non-experimental.
Narula and Guimon (2010)	Promoting the incremental upgrading of existing subsidiaries towards demand-driven R&D. R&D activity of Multi National Enterprises subsidiaries. The interaction between national innovation systems and MNEs.	Descriptive only, although paper mentions explicitly that is an IE. Non-experimental.
Thomson and Jensen (2010)	Whether subsidies and tax incentives increase R&D employment. Evaluate effectiveness of government grants	Model consists of a vertically integrated firm which produces both final goods and technology, via R&D. The firm's objective function represents the discounted stream of profit. Hamiltonian profit maximization model. Estimation with GMM-Abond. Non-experimental.
Pirtea et al. (2009)	Explore the interactions of foreign direct investments impact on the economy of host countries. Some countries separate such investments from R&D, some other have specific policies that target R&D.	Pooled panel GLS. Non-experimental.
Krammer (2009)	Study relationship between innovation's output and inputs	Poisson negative binomial regression and FGLS estimator and include various controls (year and regional dummies) to capture as much as possible of the unobserved heterogeneity.

Cerulli and Poti(2008)	Paper is a lit review of the current state of Impact Evaluations in R&D	Reviews papers and classifies them in three groups: 1. Use of Structural Models, 2. Use of non-structural models, 3. Exploiting cross-section or longitudinal data.
Ozcelik and Taymaz (2008)	Technology Development Foundation of Turkey (TTGV, in the Turkish acronym). Study the effect of direct subsidies on private R&D activity at the level of firms in the Turkish manufacturing industry.	Non-parametric matching. Firms are matched on the propensity score (the probability to receive R&D support), which is estimated by a logitmodel
Hussinger (2008)	Study the effect of public R&D subsidies on firms' private R&D investment per employee and new product sales in German manufacturing	Sample selection correction. The selection equation is estimated as a probit model on the probability of receiving public R&D funding. A tobit model is applied to test for robustness if the amount of funding is used as the endogenous variable. Newey's and Robinson's estimators are combined with the two different intercept estimators by Heckman to estimate ATET.
Aerts and Schmidt (2008)	In Germany public R&D funding relies largely on direct R&D funding; fiscal measures, like R&D tax credits, do not exist. In Flanders, accelerated depreciation for R&D capital assets and R&D tax allowances are available through the federal Belgian government.	Parametric matching.
Almeida and Teixeira (2007)	Explore asymmetric effects of patents on R&D in accordance to the level of GDP	Panel data methods (not specified which one)
Simeonova (2006)	Review of R&D policies that can be potentially evaluated in Bulgaria0	Review of existing methods to conduct surveys and data collection from firms demanding R&D support.
Aralica and Bacic (2005)	Rank Croatia's achievements in innovation policy against EU and Central and Eastern Europe countries (CEEC)	Mostly qualitative evaluation based on European Innovation Scoreboard (EIS)

Ali-Yrkkö (2005)	Public R&D funding from the Finnish Technology Agency (Tekes). Analyze how public R&D financing impacts the labor demand of companies	OLS and instrumental-variable (value of funds that are potentially awardable to firm) regressions of R&D employment on subsidies. Non-experimental.
Loof and Heshmati (2005)	Evaluates whether firms receiving public funding have on average higher R&D intensity compared to those not receiving any such support	Non-experimental matching method Nearest neighbor.
Czarnitzki and Licht (2006)	Estimate the impact of public R&D grants on firms' R&D and innovation input	Non-experimental, matching. Probit Regression for Programme Participation; All Firms. Subsample of Supported Firms vs. Firms Permanently Performing R&D. Double-diff (participation and time)
Racic et al. (2007)	To what extent and in which ways can an innovation policy facilitate innovative activities and contribute to restructuring, technological advancement and economic growth in Croatia	Non-experimental. Treatment effects of selected policy instruments and determinants of innovation activities by using a probit model. Complements with qualitative assessment using European Innovation Scoreboard
Aerts and Czarnitski (2004)	Investigate whether public R&D funding in Belgium crowds-out the private investment in the business sector.	Non-experimental matching using probit and second stage GLS.
Hanel (2004)	Government programs supporting R&D and innovation by Canadian manufacturing firms and the relationship between the support received and the R&D and innovation performance.(R&D tax credits and R&D subsidies)	3SLS-Ordered logit regressions estimating the probability that a firm uses a particular government program. two stage logit regressions estimating the probability that a firm introduces a more rather than a less original innovation. In the last section are results of ordered logit regressions estimating the impact of government programs on the share of product innovations in total sales.
David et al. (2000)	This paper explores the degree in which R&D subsidies and/or tax breaks are susbtitute of complementary to firm's outcomes. Provides a literature review and strong theoretical framework to evaluate additionality	Structural model, Endogeneity biases correction, IV, latent variables effects

	effectiveness	
Aerts (2008)	This publication reviews the types of R&D programs that can be subject to rigorous evaluation and highlights the main techniques available to assess them.	Matching methods, sample selection correction
Benavente, Crespi, and Maffioli (2007)	Does the public financing crowded out private resources? The evaluation will address the impact of the program on the beneficiaries' own financial resources devoted to R&D and innovation activities, as a test for the potential crowding out effect of the public financing	Propensity Score Matching at the Firm level. Differences in Differences (D-in-D) to identify impacts
Hall and Maffioli (2008)	TDF effectiveness is found to depend on the financing mechanism used, on the presence of non-financial constraints, on firm-university interaction, and on the characteristics of the target beneficiaries. Four levels of potential impact were considered: R&D input additionality, behavioral additionality, increases in innovative output, and improvements in performance.	PSM at the firm level
Calderon-Madrid (2009)	Mexican government introduced a fiscal stimulus plan for businesses that invest in technological activities and whose projects were presented before CONACYT (National Council on Science and Technology).	Sample selection correction and Fixed effects (F.E.) estimators
Magro et al. (2010)	Analyze behavioral additionality as the result of a regional S&T program.	Propensity Score Matching at the Firm level. Differences in Differences (D-in-D) to identify impacts

References

Aerts, K. Czarnitski, D. Cassiman, B. Hoskens, and M. Vanhee (2007). "Research, Development and innovation in Flanders," IWT Studies No. 55.

Aerts K. and D. Czarnitzki (2004). "Using innovation survey data to evaluate R&D policy: The case of Belgium," DTEW Research Report 0439, pp. 1 - 21.

Aghion, P., G.M. Angeletos, A. Banerjee, and K. Manova (2005). "Volatility and Growth: Credit Constraints and Productivity-Enhancing Investment," NBER Working Papers 11349.

Aghion, P., P. Askenazy, N. Berman, G. Cettex, and L. Eymard (2008). "Credit Constraints and the Cyclicality of R&D Investment: Evidence from France." PSE Working Papers halshs-00586744, HAL.

Aghion, P., C. Harris, P. Howitt, and J. Vickers (2001). "Competition, Imitation and Growth with Step*-by-Step Innovation,"* Review of Economic Studies, 68(3), 467—492.

Almeida, A. and A. Teixeira (2007). "Does Patenting negatively impact on R&D investment? An international panel data assessment," INESC Porto; CEMPRE; Faculdade de Economia (FEP), Universidade do Porto

Almus, M. and D. Czarnitzki (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.

Ali-Yrkkö, J. (2005). "Impact of Public R&D Financing on Employment," Working Paper No. 39 ENEPRI

Aralica, Z. and K. Bacic (2005). "Evaluation of Croatian Innovation Policy." ZIE, WP 0-21

Arrow, K. J. (1962). "Economic Welfare and the Allocation of Resources to Invention," in R.R. Nelson (ed.), The Rate and Direction of Economic Activity, Princeton University Press, N.Y.

Baghana, R. (2010). "Public R&D Subsidies and productivity: Evidence from firm-level data in Quebec," UNU-MERIT WPS #2010-055

Barlevy, G. (2007). "On the Cyclicality of Research and Development," American Economic Review, American Economic Association 97(4): 1131–1164.

Bayona-Sáez, C. and T. Garcia-Marco (2010). "Assessing the effectiveness of the Eureka Program," Research Policy. Volume 39, Issue 10, December 2010, Pages 1375–1386.

Begg, C. B., and J. A. Berlin. (1988). "Publication bias: a problem in interpreting medical data." Journal of the Royal Statistical Society, Series A 151: 419-463.

Benavente J. M., G. Crespi, and A. Maffioli (2007). "Public Support to Firm Innovation: The Chilean FONTEC Experience," OVE Working Papers 0407, Inter-American Development Bank, Office of Evaluation and Oversight (OVE).

Bloom, N., R. Griffith, and J. Van Reenen (2002). "Do R&D tax credits work? Evidence from a panel of countries 1979-1997," *Journal of Public Economics, Elsevier*, vol. 85(1)

Brown, J.D., J.S. Earle, H. Vakhitova, and V. Zheka (2010). "Innovation, Adoption, Ownership, and Productivity: Evidence from Ukraine," ESCIRRU Working Paper No. 18. Berlin, Germany: ESCIRRU (Economic and Social Consequences of Industrial Restructuring in Russia and Ukraine).

Bruno, R.L. and C. Campos (2011). "*A Systematic Review of the Effect of FDI on Economic Growth in Low Income Countries: A Meta-Regression-Analysis*," DFID: London, United Kingdom. http://discovery.ucl.ac.uk/1364958/

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

Buisseret, T. J., Cameron, H. M., and Georghiou, L. (1995). "What Difference Does It Make - Additionality in the Public Support of R-and-D in Large Firms," International Journal of Technology Management, 10(4-6), 587-600.

Calderon-Madrid, A. (2009). "Evaluación del Programa de Estímulos Fiscales al Gasto en Investigación y Desarrollo de Tecnología de las Empresas Privadas en México (EFIDT)," CONACYT. http://www.conacyt.gob.mx/registros/sinecyt/Documents/Evaluacion_del_Impacto_del_Programa_de_Estimulos_Fiscales_al_Gasto_en_IyDT.PDF

Cerulli G. and B. Potì (2008). "Evaluating the Effect of Public Subsidies on Firm R&D Activity: An Application to Italy Using the Community Innovation Survey," Ceris-CNR Working Paper, N. 09/08.

Cerulli G. and B. Potì (2010). "The differential impact of privately and publicly funded R&D on R&D investment and innovation: the Italian case," Working Papers 10, Doctoral School of Economics, Sapienza University of Rome.

Costa-Font, J., Gammill, M. and Rubert, G. (2011). "Biases in the healthcare luxury good hypothesis: A meta-regression analysis," Journal of the Royal Statistical Society A. 174.

Costa-Font,J., F. De-Alburquerque, and C. Doucouliagos (2011). "How Significant are Fiscal Interactions in Federations? A Meta-Regression Analysis," CESIFO WORKING PAPER NO. 3517. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1891839

Czarnitzki, D., P. Hanel, and J.M. Rosa (2004). "Evaluating the Impact of R&D Tax Credits on Innovation: A Microeconometric Study on Canadian Firms," ZEW Discussion Papers 04-77, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Czarnitzki, D. and A. Fier (2002). "Do innovation subsidies crowd out private investment? Evidence from the German Service Sector," Applied Economics Quarterly (Konjunkturpolitik) 48(1), 1-25.

David, P., B. Hall, and A. Toole (2000). "*Is public R&D a complement or substitute for private R&D? A review of the econometric evidence,*" Research Policy, Volume 29, Issues 4—5, 497—529.

David, P.A. and B.H. Hall (1999). "Heart of darkness, public–private interactions inside the R&D black box," Economics Discussion Paper No. 1999-W-16, March, Nuffield College, Oxford, forthcoming with revisions in Research Policy.

Doucouliagos, H. and M. Ulubasoglu (2010). "*Democracy and Economic Growth: A meta-analysis,*" Deakin University, Australia. http://www.international.ucla.edu/cms/files/Doucouliagos.pdf

Doucouliagos, C. and T.D. Stanley (2008). "Theory Competition and Selectivity: Are all economic facts greatly exaggerated?" Deakin University, Economics Working Paper, No. 2008_14.

Duguet, E. (2006). "Innovation Height, Spillovers and TFP Growth at the Firm Level: Evidence from French Manufacturing for Company Performance," Economics of Innovation and New Technology. 15 (4/5).

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.

Eesley, C.E. (2010). "Institutions and Innovation: A Literature Review of the Impact of Public R&D and Financial Institutions on Firm Innovation," SSRN 146. Stanford University.

Egger, M., G.D. Smith, M. Schneider, and C. Minder (1997). "Bias in meta-analysis detected by a simple, graphical test," British Medical Journal;315:629-634.

Erden, I.T. (2010). "A case study of Impact Analysis: Tubitak Research Support Programmes," PhD. Thesis. Middle East Technical University, Turkey.

Eurostat (2011). "Europe in Figures." Eurostat Yearbook 2011. Statistical Books. Available at: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-CD-11-001/EN/KS-CD-11-001-EN.PDF

Falk, R. (2007). "Measuring the effects of public support schemes on firms' innovation activities: Survey evidence from Austria," Research Policy, 36(5), pp: 665-79.

Fisher, R., W. Polt, and N. Vonotras (2009). "The impact of publicly funded research on innovation An analysis of European Framework Programmes for Research and Development." European Commission on Enterprise and Industry. PRO INNO Europe paper N°7.

http://www.eurosfaire.prd.fr/7pc/doc/1264491592_impact_public_research_innovation.pdf

Garcia, A., and P. A. Mohnen (2010). "Impact of government support on R&D and innovation". UNU-MERIT WPS #2010-034.

Georghiou, L., and B. Clarysse (2006). "*Introduction and synthesis*. In Government R&D Funding and Company Behaviour." Measuring Behavioural Additionality. Paris: OECD.

Glass, G. V., B. McGaw, and M.L. Smith (1981). "Meta-analysis in social research." Beverly Hills, CA: SAGE.

González, X., J. Jaumandreu, and C. Pazó (2005). "Barriers to innovation and subsidy effectiveness," *RAND Journal of Economics* 36(4), 930-949.

Goolsbee, A. (1998). "Does Government R&D Policy Mainly Benefit Scientists and Engineers?," NBER Working Papers 6532, National Bureau of Economic Research, Inc.

Gorg, H. and E. Strobl (2007). "The effect of R&D subsidies on private R&D. *Economica* Volume 74, Issue 294

Guellec, D. and B. Van Pottelsberghe (2001). "R&D and Productivity Growth: Panel Data Analysis of 16 OECD countries." OECD Economic Studies No. 33

Hall, B. (2011). "Innovation and productivity." UNU-MERIT UNU028.

Hall, B. and A. Maffiolli (2008). "Evaluating the Impact of Technology Development Funds in Emerging Economies: Evidence from Latin America." Inter- American Development Bank. OVE Doc047. Washington D.C.

Harbord R.M., M. Egger, and J.A.C. Sterne (2006). "A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints." Statistics in Medicine.

Hausman, J ., B. Hall, and Z. Griliches (1984). "*Econometric Models for Count Data with an Application to the Patents — R&D Relationship*." Econometrica, 52(4), 903—938.

Heckman, J. J., H. Ichimura, and P. Todd (1997). "Matching as an econometric evaluation estimator: evidence from evaluating a job training program," Review of Economic Studies 64(4), 605-654.

Hu, A. and G.H. Jefferson (2009). "A great wall of patents: What is behind China's recent patent explosion?," *Journal of Development Economics* No. 90

Hussinger, K. (2008). "R&D and Subsidies at the Firm Level: An Application of Parametric and Semi-Parametric Two-Step Selection Models," Journal of Applied Econometrics 23, 729-747

Imbens, G. W. and J.M. Wooldridge (2009). "Recent Developments in the Econometrics of Program Evaluation," Journal of Economic Literature 47, 5-86.

Kilponen J. and T. Santavirta (2007). "*When do R&D subsidies boost innovation? Revisiting the inverted U-shape*," Bank of Finland Disc. Paper 10.

Jefferson, G., B. Huamao, G. Xiaojing, and Y. Xiaoyun (2006)."R&D Performance in Chinese industry," Economics of Innovation and New Technology, Taylor and Francis Journals, vol. 15(4-5), pages 345-366.

Kerr, W.R. and R. Nanda (2008). "Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship," HBS Finance Working Paper No. 07-033; US Census Bureau Center for Economic Studies Paper No. CES-WP-07-33; Harvard Business School Entrepreneurial Management Working Paper No. 07-033. Available at SSRN: http://ssrn.com/abstract=999985.

Krammer, S. (2009). "Drivers of national innovation in transition: Evidence from a panel of Eastern European countries," Research Policy, Elsevier, vol. 38(5), pages 845-860, June.

Klette, T.J., J. Moen, and Z. Griliches (2000). "Do subsidies to commercial R&D reduce market failures? Microeconometric evaluation studies," Research Policy, Elsevier, vol. 29(4-5), pages 471-495, April.

Lach, S. (2002). "Do R&D subsidies stimulate or displace private R&D? Evidence from Israel," Journal of Industrial Economics 50(4), 369-390.

Lengyel, B., I. Ikawashi, and M. Szanyi (2008). "Industry Cluster and Regional Economic Growth: Evidence from Hungary." Hitotsubashi Journal of Economics Vol. 51, no. 2

Levin, R., A. Klevorick, R. Nelson, and S. Winter (1987). "Appropriating the returns from industrial R&D." Brookings Papers Econ. Activity 3 783-831.

Light, R. and D. Pillemer (1984). "Summing Up: The Science of Reviewing Research." Harvard University Press.

Heshmati, A. and H. Loof (2005). "The Impact of Public Funds on Private R&D Investments: New Evidence from a Firm Level Innovation Study," Discussion Papers 11862, MTT Agrifood Research Finland.

Lentile, D. and J. Mairesse (2009). "A policy to boost R&D: Does the R&D tax credit work?," EIB Papers 6/2009, European Investment Bank, Economics Department.

Lopes Bento, C. (2011) "Evaluation of public R&D policies: A cross-country comparison," European Network on Industrial Policy (EUNIP) International Workshop on Evaluating Innovation Policy: Methods and Applications. Italy.

Lopez, A., A.M. Reynoso, and M. Rossi (2010). "Impact Evaluation of a Program of Public Funding of Private Innovation Activities. An Econometric Study of FONTAR in Argentina," OVE Working Papers 0310, Inter-American Development Bank, Office of Evaluation and Oversight (OVE).

Lou, Y., P.C. Abrami, and S. d'Apollonia (2001). "Small group and individual learning with technology: A meta-analysis." Review of educational research, 71(3), 449-521.

Lukkonen T. (2000). "Additionality of EU framework programmes," *Research Policy*, 29(6)

Magro, E., J.M. Aranguren, and J. Wilson (2010). "Competitiveness policy evaluation as a transformative process: From theory to practice," *EU-SPRI Conference. ORKESTRA. Spain*.

Marino,M., P. Parrota, and D. Sala (2010). "New Perspectives on the Evaluation of Public R&D Funding," EPFL-WORKING-161988

Matsumoto, M., S. Yokota, K. Naito, and J. Itoh (2010). "Development of a model to estimate the economic impacts of R&D output of public research institutes," R&D Management, Volume 40, Issue 1, pp. 91-100.

Nelson, R. (1988). "Modeling the Connections in the Cross Section Between Technical Progress and R&D Intensity," Rand Journal of Economics (August 1988), pp. 478-485.

OECD (2012). "*Innovation in the Crisis Beyond,"* In OECD Science, Technology and Industry Outlook. Chapter 1. Available at: http://www.oecd.org/sti/sti-outlook-2012-chapter-1-innovation-in-the-crisis-and-beyond.pdf

OECD (2011). "R&D expenditure," in OECD Science, Technology and Industry Scoreboard 2011, OECD Publishing. doi: 10.1787/sti_scoreboard-2011-16-en.

OECD (2011). "Economic Policy Reforms 2011– Going for Growth," Chapter 2: Country Notes and Chapter 3: Structural Policies Indicators. Paris. http://dx.doi.org/10.1787/growth-2011-en

OECD (2006). "Government R&D Funding and Company Behaviour – Measuring Behavioural Additionality," Paris: OECD; Directorate for science, technology and industry, Comittee for scientific and technological policy.

Ozcelik, E., and E. Taymaz (2008). "R&D support programs in developing countries: The Turkish experience." Research Policy 37. Elsevier. Volume 37, Issue 2, March 2008, Pages 258scie.

Park, J. (2004). "International and Intersectoral R&D Spillovers in the OECD and East Asian Economies," Economic Inquiry (October 2004), pp. 739-757.

Pirtea, M., B. Dima, and L.R. Milos (2009). "An empirical analysis of the interlinkages between financial sector and economic growth," Annals of DAAAM for 2009 & Proceedings , Vol. 20, No. 1 (25. November 2009): pp. 343-344.

Račić, D., V. Cvijanović, and Z. Aralica (2007). "The Effects of the Corporate Governance System on the Innovation activities in Croatia," paper presented at the 7th International Conference on Enterprise in Transition, May 24-26, Bol, Croatia. http://www.fep.up.pt/conferencias/eaepe2007/Papers%20and%20abstracts_CD/Racic%20Cvijanovic%20Aralica.pdf

Rasmus Lentz, D.T.M. (2008). "An empirical model of growth through product innovation." Econometrica 76(6) 1317-1373.

Romanelli, E. (1989). "Environments and strategies of organization start-up: Effects on Early Survival," Administrative Science Quarterly 34(3) 369.

Romanelli, E. and M.L. Tushman (1986). "Inertia, environments, and strategic choice: A quasiexperimental design for comparative-longitudinal research." Management Science 32(5, Organization Design) 608-621.

Rosenbaum, P.R. and D.B. Rubin (1983). "The central role of the propensity score observational studies for causal effects," Biometrica 70, 41-55

Romer, P.M. (1986). "Increasing Returns and Long Run Growth," *Journal of Political Economy* 94(5), 1002-1037.

Roper, S. (2010). "Moving On: From Enterprise policy to innovation policy in Western Balkans," Working Paper No. 108. Centre for Small and Medium Enterprises. Warwick Business School

Sayek, S. (2009). "Foreign Direct Investment and Inflation," Southern Economic Journal, Southern Economic Association, vol. 76(2), pages 319-443, October

Simeonova, K. (2006). "Research and innovation in Bulgaria," *Science and Public Policy* Vol. 33 No. (5)

Simon, G. (2010). "Factors and Problems of Economic Growth in Hungary, Russia and Serbia," *International Problems UDK, Vol. LXII, No. 2, pp. 195-238* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1638818

Stanley, T.D. and Doucouliagos, C. (2010). "Picture This: A simple graph that reveals much ado about research," Journal of Economic Surveys Volume 24, Number 1, 170-191.

Stanley, T.D. and S.B. Jarrell (1989). "Meta-Regression Analysis: A Quantitative Method of Literature Surveys," Journal of Economic Surveys, Wiley Blackwell, vol. 3(2), pages 161-70.

Stanley, T.D. (2006). "Two-Stage Precision-Effect Estimation and Heckman Meta-Regression for Publication Selection Bias," Economics Series 2006_25, Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.

Streicher, G., A. Schibany, and N. Gretzmacher (2004). "Input Additionality Effects of R&D Subsidies in Austria: Empirical Evidence from Firm-level Panel Data," Institute of Technology and Regional Policy - Joanneum Research

Stern, J. M. and R.J. Simes (1997). "Publication bias: Evidence of delayed publication in a cohort study of clinical research projects," British Medical Journal, 1997 September 13; 315(7109): 640–645.

Sterne J.A.C. and M. Egger (2001). "Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis." *Journal of Clinical Epidemiology Vol.* 54:1046-1055.

Sveikauskas, L. (2007). "R&D and Productivity Growth," BLS Working Paper 408. Washington D.C.

Takalo, T., T. Tanayama, and O. Toivanen (2008). "Evaluating innovation policy: a structural treatment effect model of R&D subsidies," Bank of Finland Research Discussion Papers 7.

Thomson, R.K. and P.H. Jensen (2010). "The Effects of Public Subsidies on R&D Employment: Evidence from OECD Countries," Available at http://dx.doi.org/10.2139/ssrn.1740163

Varga, J. and J. Veld (2011). "A model-based analysis of the impact of Cohesion Policy expenditure 2000–06: Simulations with the QUEST III endogenous R&D model," Economic Modeling 28. Elsevier. Volume 28, Issues 1ww, January, Issues 1www.sciencedirec.

Van Pottelsberghe, B. (1997). "Issues in assessing the effect of interindustry R&D spillovers," *Economic Systems Research*, Vol. 9 No. (4).

Veryzer, R.W, (1998). "Discontinuous Innovation and the New Product Development Process," The Journal of Product Innovation Management 15 4: 304-321.

Wallsten S.J. (2000). "The effects of government-industry R&D programs on private R&D: the case Small Business Innovation Research Program," RAND Journal of Economics 31(1), 82-100

Zhao, M. (2006). "Conducting R&D in Countries with Weak Intellectual Property Rights Protection," Management Science 52(8) 1185.