GREENWICH POLITICAL ECONOMY RESEARCH CENTRE

R&D and productivity in OECD firms and industries: A hierarchical meta-regression analysis

Mehmet Ugur
University of Greenwich

Esref Trushin
University of Durham

Edna Solomon
University of Greenwich

And

Francesco Guidi
Univeristy of Greenwich

Year: 2015

No: *GPERC29*

GREENWICH PAPERS IN POLITICAL ECONOMY

Abstract: Effects of R&D investment on frim/industry productivity have been investigated widely thanks to pioneering contributions by Zvi Griliches and others in late 1970s and early 1980s. We aim to establish where the balance of the evidence lies and what factors may explain the variation in the research findings. Using 1,258 estimates from 65 primary studies and hierarchical meta-regression models, we report that the average elasticity and rate-of-return estimates are both positive, but smaller than those reported in prior narrative reviews and meta-analysis studies. We discuss the likely sources of upward bias in prior reviews, investigate the sources of heterogeneity in the evidence base, and discuss the implications for future research. Overall, this study contributes to existing knowledge by placing the elasticity and rate-of-return estimates under a critical spot light and providing empirically-verifiable explanations for the variation in the evidence base.

Key words: R&D, knowledge capital, productivity, meta-analysis

JEL Classification: D24, O30, O32, C49, C80

Acknowledgements: Research for this paper has been funded by ESRC under grant reference no ES/K004824/1. We thank our funders for their generous support. We also thank the participants of two workshops for their helpful comments: (i) Evaluation of R&D expenditures, firm survival firm growth and employment: UK Evidence workshop, University of Greenwich, 4 June 2015; and (ii) MAER-Net Colloquium 2013, 11 September 2013, University of Greenwich. The views expressed here are those of the authors, who are responsible for any errors or omissions.

Corresponding author: Mehmet Ugur, Professor of Economics and Institutions, University of Greenwich Business School, Park Row, London SE10 9LS. Email: M.Ugur@gre.ac.uk

1. Introduction

The relationship between research and development (R&D) investment and productivity has been a subject of major interest for researchers and policy makers for a long time. The pioneering work is that of Minasian (1969) and Griliches (1973) on R&D and productivity; and Terleckyj (1974) on rates of return to R&D. The empirical work has expanded significantly after Griliches (1979), who has articulated a lasting framework for the range of measurement, modeling and estimation issues encountered in empirical work.

The work tended to follow the so-called *primal approach*, which consists of a Cobb-Douglas production function augmented with R&D (knowledge) capital in addition to physical capital and labour. A smaller number of studies have adopted a *dual approach*, which draws on a system of factor demand equations and cost-function representations of technology. We synthesize the evidence from the primal approach only because the dual-approach studies are not only small in number and but also more heterogeneous in their model specifications.

Several narrative reviews of the extant literature exist. Of these, Mairesse and Sassenou (1991) and Mairesse and Mohnen (1994) review the literature on R&D and productivity at the firm and industry levels, respectively. Hall (1996) focuses on rates-of-return estimates, differentiating between private and social returns to R&D. A recent and comprehensive review by Hall et al. (2010) provides an authoritative account of the analytical, measurement and estimation issues that characterise the research field. Finally, two meta-analysis studies by Wieser (2005) and Moen and Thorsen (2013) provide meta-regression evidence on productivity and rates-of-return estimates.

We have identified a number of issues that justify a novel attempt at synthesizing the rich evidence base and explaining the sources of heterogeneity therein. First, existing narrative and quantitative reviews tend to present summary measures based on ‘representative’ or ‘preferred’ estimates and as such call for more effective use of all available information. More importantly, however, the evidence base may be contaminated with publication selection bias. To the extent that this is the case, summary measures based on 'preferred' or 'representative' estimates are inappropriate for inference.

The second issue relates to sampling bias in existing reviews, which follow a cascading approach that updates the list of studies covered in preceding reviews. This approach may allow for replication and extension as methods of verification, but the absence of explicit criteria for including or excluding primary studies may limit the representativeness of the sample and the generalizability of the offered synthesis.

The third issue relates to how existing reviews quantify the effects of moderating factors on the variation among primary study estimates. The *narrative* reviews rely on 'vote counting' for deciding whether a moderating factor (e.g., the estimation method, the measure of inputs or output, type of R&D or firms/industries, etc.) is associated with systematically larger or smaller productivity estimates. Of the meta-analysis studies, Wieser (2005) controls for a number moderating factors within a multiple meta-regression framework, but his sample consists of only 52 observations chosen from 17 primary studies.

To address these issues, we utilise 1,258 estimates from 65 studies and follow the best-practice recommendations for meta-analysis (Stanley et al., 2013). The remainder of this article is organised in three sections. Section 2 provides an overview of the analytical and empirical issues that characterise the research field. In section 3 we present the meta-analysis strategy - including the study search and selection criteria and the meta-regression methodology. In section 4 we present meta-regression estimates, based on 440 elasticity estimates in the level dimension, 468 elasticity estimates in the temporal dimension; and 350 estimates for rates of return. The sample consists of primary studies using OECD firm or industry data, and published in English between 1980 and July 2013.

We focus on OECD firm/industry studies for three reasons. First, differences in data quality are less likely to be a source of unobserved heterogeneity as the definition and collection of R&D

data in OECD countries has been harmonised substantially since 1963. Second, primary studies on EOECD firms/industries account for more than 80% of the existing evidence base. Finally, there are multiple studies per country over time and this provides scope for investigating whether productivity and rates-of-return estimates have varied over time and across OECD countries.

We report that the average elasticity and rate-of-return estimates are both positive, but smaller than those reported in prior narrative reviews and meta-analysis studies. We discuss the likely sources of upward bias in prior reviews and investigate the sources of heterogeneity in the evidence base. Our findings also indicate that there is room for innovation in future research in several areas, including the modelling and estimation of the rates of return on R&D, the measurement of the spillover pool, the relationship between R&D intensity and market power, and the separation of public and private R&D in productivity estimations.

1. Analytical and empirical dimensions of the research field

Primary studies on R&D and productivity usually draw on a Cobb-Douglas production function, augmented with knowledge (R&D) capital. Assuming perfect competition in factor markets and separability of the conventional inputs (physical capital and labour) from knowledge (R&D) capital, the production function can be stated as:

$$Y_{it} = Ae^{\lambda t}C_{it}^{\alpha}L_{it}^{\beta}K_{it}^{\gamma}e^{u_{it}} \qquad (1)$$

Here, Y is deflated output (sales or gross output or value-added); C is deflated physical capital stock; K is deflated knowledge capital; L is labour (number of employees or hours worked); λ is rate of disembodied technological change; and A is a constant. Taking natural logarithms and using lower-case letters to denote logged values, t to denote time and i to denote firm or industry, the empirical model can be written as:

$$y_{it} = \eta_i + \lambda t + \alpha c_{it} + \beta l_{it} + \gamma k_{it} + u_{it} \quad (2a)$$

The log of technical progress ($Ae^{\lambda t}$) yields a firm- or industry-specific effect (ηi) and a time effect (λt). In (2a), returns to scale are assumed as constant. However, this assumption can be relaxed and returns to scale can be tested explicitly by subtracting logged labour from both sides of (2a).

$$(y_{it} - l_{it}) = \eta_i + \lambda t + \alpha(c_{it} - l_{it}) + \gamma(k_{it} - l_{it}) + (\mu - 1)l_{it} + u_{it} \quad (2b)$$

Here, $\mu = \alpha+\beta+\gamma$ and implies constant returns to scale if $\mu = 1$; but variable returns to scale otherwise. The coefficient of interest in both (2a) and (2b) is (γ) – the *output elasticity* with respect to knowledge capital.

Usually, the R&D capital (K) is constructed with the perpetual inventory method (PIM), assuming a growth rate of 5% for R&D investment prior to initial year and a depreciation rate of 15%.[1] The consensus in the literature is that assumed rates of growth or depreciation do not alter the elasticity estimates (Hall and Mairesse, 1995; Bartelsman et al, 1996; Verspagen, 1995; Griliches and Mairesse, 1981; Harhoff, 1994; Hall and Mairesse, 1995; Bartelsman et al., 1996).[2] Therefore, we do not control either for growth or depreciation rates as potential sources of variation in the evidence base. Nevertheless, PIM's appropriateness for constructing R&D capital has been debated widely

[1] The growth rate for R&D investment can also be calculated from the R&D series over a period of τ years prior to the initial year if the R&D series is sufficiently long. The depreciation rate of 15% is informed by findings in a number of some studies, which range from 10% to 36% (Bosworth, 1978; Klette, 1994; Pakes and Schankerman, 1984; Hall, 2005).
[2] Assumed depreciation rate is not relevant when rates of return are estimated because the latter are based on R&D intensity rather than R&D capital.

(see, for example, Klette, 1991; 1996). Several contributors indicate that development of novel methods in this area constitutes promising avenues for future research (Griliches, 1979; Bitzer and Stephan, 2007; and Hall et al., 2010). Therefore, we investigate whether estimates based on other methods differ systematically from those based on PIM.

The second contentious issue is whether elasticities or rates of return should be equalised between firms/industries. Assuming elasticity equalisation overlooks the possibility that firms may choose different factor shares depending on the competitive equilibria they are faced with. Hence, a substantial number of contributors assume rates-of-return equalisation, which is more compatible with the assumption of competitive markets.

To obtain rates-of-return estimates, model (2a) is expressed in first-difference, yielding:

$$\Delta y_{it} = \Delta\lambda t + \alpha\Delta c_{it} + \beta\Delta l_{it} + \gamma\Delta k_{it} + \Delta u_{it} \tag{3a}$$

Note that the firm- or industry-specific fixed effect term (η_i) has disappeared and the time effect is now a growth-rate effect relative to initial observation rather than a level effect. Assuming that the depreciation rate (δ) is sufficiently close to zero and recalling that the elasticity of output with respect to R&D capital is given by $\gamma = (\delta Y_{it}/\delta K_i)\,(K_{it}/Y_{it}) = \rho(K_{it}/Y_{it})$, equation (3a) can be rewritten as (3b) below, where ρ is the gross rate of return on R&D investment and (R_{it}/Y_{it}) is R&D intensity.[3]

$$\Delta y_{it} = \Delta\lambda t + \alpha\Delta c_{it} + \beta\Delta l_{it} + \rho\frac{R_{it}}{Y_{it}} + \Delta u_{it} \tag{3b}$$

[3] By definition, the elasticity of output with respect to R&D capital is $\gamma = (\delta Y_{it}/\delta K_i)\,(K_{it}/Y_{it})$. Given that $(\delta Y_{it}/\delta K_i) = \rho$ is the marginal productivity of R&D capital, (3a) can be re-written as: $\Delta y_{it} = \Delta\lambda t + \alpha\Delta c_{it} + \beta\Delta l_{it} + \rho(K_{it}/Y_{it})\,\Delta k_{it} + \Delta u_{it}$. Then the term for knowledge capital simplifies as follows: $\rho(K_{it}/Y_{it})\,\Delta k_{it} = \rho(K_{it}/Y_{it})\,(\Delta K_{it}/K_{it}) = \rho(\Delta K_{it}/Y_{it}) = \rho(K_{it} - K_{it-1})/Y_{it} = \rho\frac{(1-\delta)K_{it-1}+R_{it}-K_{it-1}}{Y_{it}} = \rho\frac{R_{it}-\delta K_{it-1}}{Y_{it}} \cong \rho\frac{R_{it}}{Y_{it}}$ if rate of depreciation (δ) is close to zero.

Model (3b) allows for estimating gross rates of return on R&D investment *directly*, using output growth. Direct rates of return can be estimated also by using total factor productivity (TFP) growth by subtracting conventional inputs (physical capital and labour) from both sides of (3b):

$$\Delta TFP_{it} = \Delta\lambda t + \rho \frac{R_{it}}{Y_{it}} + \Delta u_{it} \qquad (3c)$$

Alternatively, R&D rates of return can also be obtained *indirectly*, using the definition of the R&D elasticity (γ).[4] A small number of studies report indirectly-measured rates-of-return estimates. We included such estimates in the meta-analysis only if they were reported together with their standard errors.

Differences in econometric specification are significant sources of variation in the evidence base. For example, the variables may be expressed in levels or first-difference; the production function may or may not be augmented with a measure of spillovers or with time-industry dummies; and different estimators may be used.

To highlight the difference between elasticity estimates based on level and first-differenced data, consider the total error in 2a or 2b, which is: $\eta_i + \lambda_t + u_{it}$. Estimating 2a or 2b with level data is feasible if one assumes η_i and λ_t are constant across all units and time periods, respectively. Another approach would be to maintain that assumption about fixed-effects (η_i) but eliminate the time effects by estimating the model for each period or by averaging over a time period. In all of these cases, the elasticity estimates (γ) are in the *level dimension* as they reflect cross-sectional variation in the *levels* of R&D capital and other inputs.

[4] As indicated above, $\gamma = (\delta Y_{it}/\delta K_i)\,(K_{it}/Y_{it}) = \rho(K_{it}/Y_{it})$, where ρ is the marginal product of R&D capital. Hence, $\rho = \gamma/(K_{it}/Y_{it}) = \gamma(Y_{it}/K_{it})$, where Y_{it} and K_{it} are sample means of output and R&D capital, respectively.

Alternatively, (2a) or (2b) can be estimated by first-differencing or by using a within estimator that utilises deviations from the mean. In both cases, the unit-specific fixed effect (η_i) disappears and the elasticity estimates are referred to as *elasticities in the temporal dimension*.

Elasticity estimates in both dimensions will be consistent and similar if (2a) and (2b) are specified correctly and the variables are free of measurement errors. In practice, however, the model is estimated with different control variables; and measurement errors cannot be ruled out. Therefore, and in accordance with the existing practice, we analyse the elasticity estimates in the *level* and *temporal* dimensions separately.

Unlike elasticities, rates of return are difficult to interpret for two reasons. When estimated *indirectly*, they could be interpreted either as a risk premium or a supra-normal rate of profit on R&D investments (see, Griliches, 1979; Schankerman, 1981; and Hall et al., 2010). However, Griliches (1980a: 389) points out that this interpretation is valid only when the elasticity estimate (γ) used to derive them is in the level (as opposed to temporal) dimension. Secondly, the rates of return reported in primary studies (estimated directly or indirectly) measure private returns only, which may be much smaller than social returns in the presence of externalities (spillovers). In this meta-analysis, we analyse the gross private rates of return only, but we control for whether the latter differ between studies depending on whether spillovers are controlled for as an additional source of productivity.

Model specification and estimation methods are additional sources of heterogeneity in the evidence base. For example, some studies control for spill-over effects (Aiello and Cardamone, 2005: Blanchard et al., 2005; Cincera, 1998; Hanel, 2000; Los and Verspagen, 2000, etc.); whilst others do not (Bartelsman et al., 1996; Griliches, 1980a and 1980b; Griliches and Lichtenberg, 1984; and Hall, 1993). Secondly, most of the primary studies use a standard Cobb-Douglas production function but a minority utilizes a *translog* version of the latter (e.g., Cameron et al., 2005; Lehto, 2007; and Smith et al., 2004). Third, some studies control for endogeneity via instrumented variable techniques including the general method of moments (GMM) (Aldieri et al, 2008; Ballot et al, 2006; Griffith et al., 2004); whilst other do not. Finally, some studies adopt panel-cointegration and cross-sectional dependence models (which call enhanced panel-data models) (e.g., Anon and Higon, 2007; Eberhardt et al, 2013). We control for such sources of observable heterogeneity by coding each estimates with the relevant model specification and estimation methods.

We also control for differences in measurement. One measurement issue is *double-counting*, which arises when R&D capital expenditures and R&D personnel are counted twice: on their own and as part of the physical capital (C) and labour (L). It is often reported that failure to correct for double counting leads to downward bias in the elasticity estimates (Griliches, 1979, 1980a; Harhoff, 1994; Mairesse and Hall, 1994; Hall et al., 2010); and in rates-of-return estimated indirectly (Schankerman, 1981). Therefore, we control for whether primary studies correct for double counting in both elasticity and rates-of-return estimations. Another measurement issue relates to output. Cunéo and Mairesse (1984) and Mairesse and Hall (1994) report that elasticity estimates based on value-added do not differ from those based on sales without including materials as an additional input. However, Griliches and Mairesse (1981) indicate that elasticity estimates based on value added tend to be smaller than those based on sales without materials. Therefore, we control for measurement of output as a potential source of heterogeneity in the evidence base.

The final set of moderating factors we control for include publication type (whether the primary study is a journal article or a working paper against a reference category consisting of all others); countries (France, Germany, UK or US versus other OECD countries); time period (whether the median of the data period is before or after 1980); R&D type (private *versus* public R&D); R&D intensity; and whether the underlying data is at the firm or industry level. Definitions of and summary statistics for all moderating variables are given in Tables A1 and A2 in the *Appendix*.

2. **Meta-analysis: procedures and method**

We use meta-regression analysis (MRA) to provide verifiable estimates for: (a) the 'average' productivity or rate-of-return estimate after taking account of publication selection bias; (b) the extent of publication selection bias; and (c) the effects of a wide range of moderating factors on the variation in the evidence base. We follow the best-practice recommendations in Stanley et al. (2013) and identify the relevant studies by searching in nine (9) databases, using 13 search terms for *Title* and 20 search terms for *Abstract* fields (*Table P1*).[5] We also use the snowballing approach and identify 32 studies through backward citations.

In the first stage, two reviewers read the title and abstract information of 979 total hits to select the relevant studies. In stage two, we made inclusion and exclusion decisions based on full-text information. In both stages, all studies have been coded with de-selection and exclusion criteria specified in *Table P2* (see note 5). We have de-selected 343 studies on the basis of relevance criteria and 297 studies as duplicates. In stage two, we excluded 274 studies by invoking at least one of the exclusion criteria. The final sample consists of 65 primary studies that report elasticity and private rates-of-return estimates; and published in English between 1980 and July 2013. The frequencies with which a de-selection criterion is invoked in stage one are given in *Table P3*; and those for exclusion decisions in stage two are given in *Table P4* (see note 5). Data extraction yielded 1,262 estimates; but 4 of these have been excluded from estimations as they are found to have undue influence.[6] Hence, the meta-analysis is based on 1,258 estimates, of which 440 are elasticity estimates in the level dimension, 468 are elasticity estimates in the temporal dimension, and 350 are rate-of-return estimates.

[5] The search, inclusion/exclusion and critical evaluation criteria have been specified in a protocol, the link to which will be provided after the completion of the anonymous review process. Table P1 and all other tables with a 'P' prefix are in the said protocol.
[6] Observations with undue influence are identified through the DFBETA routine in *Stata*, which indicates undue influence if |DFBETA|>1.

First, we calculate fixed-effect weighted means (FEWMs) per study and for each evidence pool, in accordance with:

$$\bar{X}_{fee} = \sum e_i(1/SE_i^2)/\sum(1/SE_i^2) \quad (6)$$

Here, e_i is the *elasticity* or *rate-of-return* estimate and $1/SE_i^2$ is precision-squared, which assigns lower weights to estimates with larger standard errors. FEWMs are more reliable than simple means, but they may conceal a high degree of heterogeneity among the estimates extracted from each study.

To estimate 'genuine effect' beyond publication selection bias, we draw on meta-regression analysis (MRA) models proposed by Stanley (2005, 2008), Doucouliagos and Stanley (2012, 2013) and Stanley and Doucouliagos (2012, 2013a). The underpinning theoretical framework is that of Egger et al. (1997), who postulate that researchers with small samples and large standard errors would search intensely across model specifications, econometric techniques and data measures to find sufficiently large (hence statistically significant) effect-size estimates. Hence, denoting effect size (i.e., the elasticity or rates-of-return estimate) with e_i and the standard error with SEi:

$$e_i = \beta + \alpha SE_i + u_i \quad (7)$$

Rejecting the null hypothesis of $\alpha = 0$ indicates the presence of publication selection bias. This is in line with increasing emphasis on the need to control for selection bias in both social sciences and medical research (see, Card and Krueger, 1995; Dickersin and Min, 1993; Ioannidis, 2005; and Simmons et al., 2011). The test is also known as funnel-asymmetry test (FAT), which evaluates the asymmetry of the funnel graphs that chart the effect-size estimates against their precisions.[7] Testing for $\beta = 0$ is a test for whether genuine effect exists beyond selection bias.

However, estimating (7) poses several issues. First, the model is heteroskedastic because the effect-size estimates have widely different standard errors (hence variances) that violate the assumption of independently and identically distributed (i.i.d.) error term (u_i). To address this issue, Stanley (2008) and Stanley and Doucouliagos (2012) propose a weighted least squares (WLS) version, obtained by dividing both sides of (7) with precision ($1/SE_i$), leading to:

$$t_i = \beta \left({}^{1}/_{SE_i} \right) + \alpha + v_i \qquad (8)$$

Here t_i is t-values reported in or calculated from primary studies; and the error term $v_i = u_i/SE_i$. Under the Gauss-Markov theorem, OLS estimation of (8) yields minimum-variance linear unbiased estimates. Testing for $\alpha = 0$ is a test for publication selection bias whereas testing for $\beta = 0$ is a 'genuine effect' test (or precision-effect test - PET) after controlling for selection bias. The selection bias is considered as *substantial* if $|\alpha| \geq 1$; and as *severe* if $|\alpha| \geq 2$ (Doucouliagos and Stanley, 2009; 2012).

The second issue is whether there is random variation between and within studies beyond idiosyncratic errors. We address this issue by estimating a hierarchical version in which primary-study estimates (the lower-level observations) are nested within primary studies (higher-level clusters). Hence:

$$t_{ij} = \alpha + \beta \left({}^{1}/_{SE_{ij}} \right) + v_j + \varepsilon_{ij} \qquad (9)$$

[7] There is a mistaken presumption that the model proposed by Egger et al. (1997) makes the detection of publication selection bias almost inevitable because of the positive association between effect-size estimates and their standard errors (or because of the negative association between effect-size estimates and their precision). On the contrary, simulation results in Stanley (2008) indicate that the funnel asymmetry test based on Egger et al (1997) has low power – i.e., the test tends to fail in detecting publication selection when the latter actually exists.

Here, subscripts j and i refer to higher-level clusters and lower-level observations, respectively; and ε_{ij} is a multivariate-normal error term with mean zero and variance matrix σ^2R, with R containing the residual-variance parameters. The study-level random effect (v_j) is assumed orthogonal to the error term ε_{ij}. The random effects (v_j) are not estimated directly, but their variance (or standard error) is.

Hierarchical models have two advantages over standard linear models. First, they allow for inclusion of random deviations other than those associated with the idiosyncratic error term. Secondly, they allow for modeling the random deviations as both between- and within-study variations (Demidenko, 2004; McCulloch et al., 2008). As such, hierarchical models are particularly relevant for meta-analysis of the research evidence, which reflects a high level of heterogeneity across and within primary studies.

To account for *both* between- and within-study variation, the random-effect term in (9) can be modelled as follows:

$$t_{ij} = \alpha + \beta\left({}^{1}\!/_{SE_{ij}}\right) + v_{0j} + v_{1j}\left({}^{1}\!/_{SE_{ij}}\right) + \varepsilon_{ij} \qquad (10)$$

Here, v_{0j} captures the between-study variation (so-called random intercepts) and v_{1j} captures within-study variation between slopes.

The third issue is that Egger et al. (1997) assume a linear relationship between primary-study estimates and their standard errors. However, Moreno et al (2009) and Stanley and Doucouliagos (2014) provide simulation evidence indicating that a quadratic specification is superior if ‘genuine effect’ exists beyond selection bias – i.e., if the PET in (9) or (10) rejects the null hypothesis of zero effect. Then, the correct specification is referred to as precision-effect test corrected for standard errors (PEESE) and can be stated as follows:

$$t_{ij} = \alpha SE_{ij} + \beta\left({}^{1}\!/_{SE_{ij}}\right) + v_j + \varepsilon_{ij} \qquad \text{(Random intercepts only)}$$

(11)

$$t_{ij} = \alpha SE_{ij} + \beta\left({}^{1}/_{SE_{ij}}\right) + v_{0j} + v_{1j}\left({}^{1}/_{SE_{ij}}\right) + \varepsilon_{ij} \quad \text{(Random intercepts and slopes)}$$

We select the appropriate model on the basis of likelihood ratio (LR) tests, with the null hypothesis that the comparison model is nested within the preferred model. A rejection of the null hypothesis indicates that the preferred model is a better fit for the data at hand. The testing procedure can be summarised as follows: (i) estimate model (9) and establish whether the LR test justifies model (9) as opposed to standard OLS (8); (ii) if (9) is preferred, test for random intercepts-only model (9) *versus* the model with random intercepts and slopes (10); and (iii) if the precision-effect test (PET) in (ii) confirms presence of non-zero effect, conduct LR test to choose between random-intercepts or random-intercepts-and-slopes version of the PEESE model in (11).

We also address the issue of overly influential observations, using the DFBETA routine in *Stata*. This involves calculating the difference in the regression coefficient when the i^{th} observation is included and excluded. The difference is scaled by the estimated standard error of the coefficient; and observations with |DFBETA| > 1 are excluded from the estimation.

The effect-size estimate (β) obtained in (11) is an 'average' R&D elasticity or rate of return, taking account of publication selection bias and the quadratic relationship between primary-study estimates and their standard errors. Although this is more reliable than simple or fixed-effect weighted means, its out-of-sample generalizability may be limited due to excessive heterogeneity in the evidence base. To identify the sources of heterogeneity, we estimate a multivariate meta-regression model in which we control for a wide range of moderating factors (i.e., observable sources of heterogeneity) that reflect the dimensions of the research field. This can be stated as follows:

$$t_i = \alpha + \beta(1/SE_{ij}) + \textstyle\sum_k \theta_k Z_k(1/SE_{ij}) + v_j + \varepsilon_{ij} \quad \text{(Random intercepts only)} \qquad (12)$$

$$t_i = \alpha + \beta(1/SE_{ij}) + \sum_k \theta_k Z_k(1/SE_{ij}) + v_{0j} + v_{1j}\left(1/_{SE_{ij}}\right) + \varepsilon_{ij}$$

(Random intercepts and slopes)

All terms and subscripts are as defined above. The *kx1* vector of moderating variables (Z_k) reflects the dimensions of the research field and constitutes the observable sources of heterogeneity in the evidence base. Because the inclination towards publishing statistically-significant findings is pervasive in social sciences, medicine and physical sciences in general (Card and Krueger, 1995; Dickersin and Min, 1993; Ioannidis, 2005; and Simmons et al., 2011), it is important that the Z-variables are interacted with precision.

To minimise the risk of multicollinearity and over-fitting, we estimate (12) through a general-to-specific estimation routine, whereby we omit the *most insignificant* variables one at a time until all remaining covariates are statistically significant. We present the findings from the specific and general models side by side to: (a) establish the extent of congruence between the significant moderating factors; and (ii) identify the range of moderating variables that do not affect the variation in the evidence base.

3. **Meta-analysis results: R&D effects and sources of heterogeneity**

We report three sets of evidence on R&D elasticities and rates of return: (1) fixed-effect weighted means (FEWMs); (2) 'average' effect-size estimates that take account of publication selection bias; and (3) multivariate meta-regression evidence on how moderating variables affect the estimates reported in primary studies.

3.1 Fixed-effect weighted means (FEWMs)

Table 1 presents FEWMs for elasticity estimates in the level and temporal dimensions (1a and 1b) and for rates-of-return estimates (1c). The FEWM is 0.053 for the sample of elasticities in the level dimension; 0.012 for the sample of elasticities in the temporal dimension; and 11.5% for the rates-of-return sample. They indicate that knowledge capital has positive effects on productivity, but the effects are smaller than what is reported in existing reviews. For example Wieser (2005) report an average elasticity in excess of 0.10. So does Hall et al. (2010), who report a typical elasticity of 0.10 or larger in the level dimension and 0.08 when the level and temporal dimensions are taken together. Typical mean values for rates of return are reported at 20%-30% by Hall et al (2010); at 28.3% on average by Wieser (2005); and at 18.2% on average in Moen and Thorsen (2013).

Study-based FEWMs also indicate a high degree of heterogeneity in the research field. Across studies, they range from -0.262 to 0.648 for elasticities in the level dimension; from -0.328 to 0.810 for elasticities in the temporal dimension; and from -73.7% to 231% for rates of return. The extent of heterogeneity is high even within studies, with the highest levels of variation ranging: from -0.107 to 0.462 for elasticities in the level dimension in Eberhardt et al (2013); from -0.304 to 0.251 for elasticities in the temporal dimension in O'Mahoney and Vecchi (2009); and from 19% to 231% for rates of return in Link (1981).

Table 1a: Fixed-effect weighted means (FEWMs) for elasticities in the level dimension

Study	Pbn. type	Unit of analysis	Country	Observations	FEWM	Std. Dev.	Min	Max
1. Aldieri et al (2008)	Journal article	Firm	US	4	0.271	0.018	0.250	0.290
2. Ballot et al (2006)	Journal article	Firm	OECD-other, France	10	0.054	0.012	0.025	0.135
3. Bartelsman (1990)	Working paper	Firm	US	6	0.006	0.007	-0.005	0.149
4. Bartelsman et al (1996)	Report	Firm	OECD-other, France	12	0.016	0.013	0.003	0.076
5. Blanchard et al (2005)	Journal article	Firm	France	6	0.085	0.018	0.080	0.168
6. Boler et al (2012)	Working paper	Firm	OECD-other, France	5	0.034	0.021	0.020	0.100
7. Bond et al (2002)	Working paper	Firm	UK, Germany	6	0.061	0.010	0.053	0.083
8. Bonte (2003)	Journal article	Industry	Germany	2	0.026	0.002	0.024	0.028
9. Cincera (1998)	Thesis	Firm	OECD-other, France	10	0.136	0.070	0.080	0.470
10. Cuneo and Mairesse (1984)	Working paper	Firm	France	10	0.159	0.060	0.058	0.209
11. Eberhardt et al (2013)	Journal article	Industry	OECD-other, France	15	0.092	0.023	-0.107	0.462
12. Frantzen (2002)	Journal article	Industry	OECD-mixed	7	0.164	0.022	0.147	0.202
13. Griffith et al (2006)	Journal article	Firm	UK	14	0.020	0.008	0.004	0.033
14. Griliches (1980b)	Book chapter	Firm	US	22	0.059	0.019	0.029	0.186
15. Griliches (1998)	Journal article	Firm	US	12	0.122	0.028	0.044	0.247
16. Griliches and Mairesse (1981)	Working paper	Firm	US	14	0.146	0.082	-0.007	0.292
17. Hall (1993)	Journal article	Firm	US	75	0.028	0.029	-0.262	0.648
18. Hall and Mairesse (1995)	Journal article	Firm	US	14	0.230	0.028	0.176	0.254
19. Harhoff (1994)	Working paper	Firm	Germany	13	0.136	0.019	0.090	0.163
20. Hsing and Lin (1998)	Journal article	Firm	US	2	0.204	0.000	0.204	0.204
21. Kafourous (2005)	Journal article	Firm	UK	17	0.038	0.041	-0.091	0.152
22. Kwon and Inui (2003)	Journal article	Firm	OECD-mixed	22	0.101	0.018	0.071	0.130
23. Lehto (2007)	Journal article	Firm	OECD-mixed	13	0.033	0.011	0.014	0.059
24. Mairesse and Hall (1996)	Working paper	Firm	France, US	29	0.047	0.047	-0.193	0.246
25. Ortega-Argiles et al (2010)	Journal article	Firm, Industry	OECD-other	8	0.082	0.038	0.017	0.169
26. O'Mahoney and Vecchi (2000)	Book chapter	Firm	OECD-other	1	0.098	n.a.	0.098	0.098
27. Rogers (2010)	Journal article	Firm	UK	12	0.012	0.019	0.009	0.238
28. Schankerman (1981)	Journal article	Firm	US	18	0.069	0.047	0.018	0.292

29. Smith et al (2004)	Journal article	Firm	OECD-other	8	0.090	0.011	0.080	0.125
30. Verspagen (1995)	Journal article	Industry	France, Germany, UK, OECD-other	55	0.022	0.031	-0.024	0.171
Overall	**All**	**All**	**All**	**443**	**0.053**	**0.055**	**-0.262**	**0.648**

Table 1b: Fixed-effect weighted means (FEWMs) for elasticities in the temporal dimension

Study	Publication type	Unit of analysis	Country	Observations	FEWM	Std. Dev.	Min	Max
1. Aiello and Cardamone (2005)	Journal article	Firm	OECD-other	4	0.055	0.004	0.053	0.090
2. Aldieri et al (2008)	Journal article	Firm	OECD-other, US	12	0.170	0.072	0.090	0.460
3. Anon and Higon (2007)	Journal article	Industry	UK	4	0.307	0.022	0.281	0.331
4. Bartelsman (1990)	Working paper	Firm	US	6	0.033	0.067	-0.005	0.180
5. Bartelsman et al (1996)	Report	Firm	OECD-other	10	0.071	0.045	0.028	0.247
6. Blanchard et al (2005)	Journal	Firm	France	1	0.013	.	0.013	0.013
7. Bond et al (2002)	Working paper	Firm	Germany, UK	6	0.024	0.079	-0.328	0.053
8. Bonte (2003)	Journal article	Industry	Germany	6	0.007	0.002	0.006	0.016
9. Branstetter (1996)	Working paper	Firm	OECD-other, US	2	0.056	0.115	0.013	0.360
10. Cincera (1998)	Thesis	Firm	OECD-other	48	0.192	0.062	0.040	0.480
11. Cuneo and Mairesse (1984)	Working paper	Firm	France	10	0.106	0.061	0.027	0.229
12. Doraszelski and Jaumandreu (2013)	Journal article	Firm	OECD-other	18	0.014	0.015	-0.003	0.075
13. Eberhardt et al (2013)	Journal article	Industry	OECD-other	3	0.053	0.015	0.024	0.063
14. Goto and Suzuki (1989)	Journal article	Industry	OECD-other	21	0.334	0.147	0.190	0.810
15. Griliches (1980a)	Journal article	Industry	US	3	0.050	0.015	0.026	0.067
16. Griliches (1980b)	Book chapter	Firm	US	37	0.073	0.021	0.011	0.232
17. Griliches (1998)	Journal article	Firm	US	5	0.108	0.009	0.095	0.110
18. Griliches and Mairesse (1981)	Working paper	Firm	US	18	0.093	0.079	-0.062	0.270
19. Griliches and Mairesse (1991b)	Book chapter	Firm	OECD-other	2	0.025	0.005	0.020	0.030
20. Hall (1993)	Journal article	Firm	US	10	0.023	0.019	-0.011	0.175
21. Hall and Mairesse (1995)	Journal article	Firm	US	42	0.072	0.057	-0.001	0.320
22. Harhoff (1994)	Working paper	Firm	Germany	46	0.113	0.061	-0.072	0.258
23. Harhoff (2000)	Journal article	Firm	Germany	5	0.068	0.001	0.067	0.069
24. Kwon and Inui (2003)	Working paper	Firm	OECD-mixed	60	0.046	0.038	-0.010	0.149
25. Lehto (2007)	Journal article	Firm	OECD-mixed	5	0.023	0.012	0.003	0.035

26. Los and Verspagen (2000)	Journal article	Firm	US	12	-0.001	0.004	-0.008	0.102
27. Mairesse and Hall (1996)	Working paper	Firm	US, France	34	0.036	0.059	-0.132	0.176
28. Ortega-Argiles et al (2010)	Journal article	Firm, Industry	OECD-other	8	0.041	0.099	-0.120	0.234
29. O'Mahoney and Vecchi (2000)	Book chapter	Firm	OECD-other, US	8	0.266	0.067	0.042	0.354
30. O'Mahoney and Vecchi (2009)	Journal article	Firm	OECD-other	9	0.149	0.116	-0.304	0.251
31. Smith et al (2004)	Journal article	Firm	OECD-other	2	0.086	0.000	0.086	0.088
32. Verspagen (1997)	Journal article	Industry	OECD-other	12	0.076	0.032	0.018	0.177
Overall	**All**	**All**	**All**	**469**	**0.012**	**0.040**	**-0.328**	**0.810**

Table 1c: Fixed-effect weighted means (FEWMs) for rate-of-return estimates

Study	Publication type	Unit of analysis	Country	Observations	FEWM	Std. Dev.	Min	Max
1. Bartelsman et al (1996)	Report	Firm	OECD-other	9	0.112	0.102	-0.004	0.348
2. Cameron et al (2005)	Journal article	Industry	UK	9	0.635	0.127	0.496	0.901
3. Cincera (1998)	Thesis	Firm	OECD-other	1	0.380	.	0.380	0.380
4. Clark and Griliches (1998)	Book chapter	Firm	US	6	0.190	0.008	0.180	0.200
5. Griffith et al (2004)	Journal article	Industry	OECD-other	15	0.479	0.095	0.343	0.857
6. Griliches (1980a)	Journal article	Industry	US	2	0.042	0.014	0.029	0.058
7. Griliches and Lichtenberg (1984)	Journal article	Firm	US	20	0.178	0.122	0.040	0.762
8. Griliches and Mairesse (1991a)	Book chapter	Firm	OECD-other, US	6	0.316	0.090	0.203	0.562
9. Griliches and Mairesse (1991b)	Book chapter	Firm	France, US	13	0.204	0.135	-0.550	0.450
10. Hall and Mairesse (1995)	Journal article	Firm	US	20	0.169	0.097	-0.013	0.341
11. Hanel (2000)	Journal article	Industry	OECD-other	8	0.168	0.080	0.077	0.338
12. Harhoff (1994)	Working paper	Firm	Germany	6	0.226	0.024	0.189	0.297
13. Heshmati and Hyesung (2011)	Journal article	Firm	OECD-other	2	0.128	0.000	0.128	0.129
14. Klette (1991)	Working paper	Firm	OECD-other	20	0.110	0.012	0.082	0.176
15. Kwon and Inui (2003)	Working paper	Firm	OECD-other	2	0.225	0.069	0.163	0.301
16. Lichtenberg and Siegel (1991)	Journal article	Firm	US	33	0.185	0.168	-0.120	1.926
17. Link (1981)	Journal article	Firm	US	2	0.252	0.360	0.190	2.310
18. Link (1983)	Journal article	Firm	OECD-other	2	0.050	0.007	0.047	0.063
19. Lokshin et al (2008)	Journal article	Firm	OECD-other	4	0.216	0.084	0.137	0.307
20. Mansfield (1980)	Journal article	Firm	US	25	0.063	0.068	-0.180	1.780
21. Mate-Garcia and Rodriguez-Fernandez (2008)	Journal article	Firm	OECD-other	1	0.266	.	0.266	0.266
22. Medda et al (2003)	Working paper	Firm	OECD-other	2	0.319	0.036	0.290	0.364

23. Odagiri (1983)	Journal article	Firm	OECD-other	2	0.185	0.217	-0.475	0.256
24. Odagiri and Iwata (1986)	Journal article	Firm	OECD-other	4	0.150	0.032	0.113	0.201
25. Rogers (2010)	Journal article	Firm	UK	18	0.144	0.064	-0.049	0.610
26. Scherer (1982)	Journal article	Industry	US	4	0.143	0.087	0.001	0.210
27. Scherer (1983)	Journal article	Industry	US	4	0.244	0.080	0.200	0.476
28. Sterlacchini (1989)	Journal article	Industry	UK	6	0.124	0.034	0.090	0.190
29. Sveikauskas (1981)	Journal article	Industry	US	21	0.082	0.077	0.039	0.394
30. Terleckyj (1980)	Book chapter	Industry	US	12	0.156	0.143	-0.180	0.370
31. van Meijl (1997)	Journal article	Industry	France	15	0.118	0.051	0.010	0.190
32. Verspagen (1995)	Journal article	Industry	Fr., Germ. UK, OECD-other	28	0.068	0.036	-0.737	0.524
33. Wakelin (2001)	Journal article	Firm	UK	14	0.269	0.139	-0.210	0.640
34. Wolff and Nadiri (1993)	Journal article	Industry	US	14	0.134	0.102	-0.087	0.612
Overall	**All**	**All**	**All**	**350**	**0.115**	**0.103**	**-0.737**	**2.310**

Given the extent of heterogeneity, it is difficult to make inferences on the basis of representative estimates chosen by primary-study authors or reviewers. The risk of incorrect inference is higher if the primary-study estimates are contaminated with publication selection bias. Therefore, and as a second step towards correct inference, we estimate the average elasticity and rate-of-return estimates after controlling for publication selection bias.

4.2 Elasticities and rates of return beyond selection bias

PET/FAT and PEESE results from bivariate hierarchical meta-regressions are given in Table 2. The models are fitted with random intercepts and random slopes in accordance with LR test results. Standard deviations of the random slopes and the residuals are all significant - indicating that the hierarchical model specification is preferable due to presence of between- and within-study variations that cannot be explained by sampling differences.

Table 2: PET/FAT and PEESE estimates of average elasticities and rates of return

	PANEL A – PET/FAT RESULTS			PANEL B – PEESE RESULTS		
Dependent variable: t-value	**Elasticities (Level)**	**Elasticities (Temporal)**	**Rates of return**	**Elasticities (Level)**	**Elasticities (Temporal)**	**Rates of returns**
Precision	0.067***	0.046***	0.072***	0.079***	0.057***	0.117***
	(0.012)	(0.015)	(0.012)	(0.011)	(0.012)	(0.014)
Constant	1.368***	1.274***	1.386***			
	(0.453)	(0.344)	(0.199)			
Standard Error				2.692	2.035***	0.108
				(6.933)	(0.641)	(0.337)
Random-effects						
Random slopes	0.053***	0.049***	0.027***	0.053***	0.052***	0.049***
Random intercepts	0.932	0.000	0.491***	1.131	0.757	1.100
Residual	4.108***	3.877***	1.713**	4.141***	3.811***	1.706**
Model diagnostics						
Observations	440	468	350	440	468	350
Number of studies (clusters)	30	32	34	30	32	34
Log likelihood (HM)	-1288	-1315	-701	-1292	-1321	-718
Chi-square	34	10	34	52	52	66
P > Chi-square	0	0	0	0	0	0
Log likelihood (OLS)	-1494	-1382	-720	-1494	-1407	-763

Notes: *, **, *** indicate significance at 10%, 5% and 1*, respectively. All models are estimated with random intercepts and random slopes, in accordance with LR tests. Significance of random-effect terms is based on standard errors (not reported here) for the natural logarithms of the standard deviations. Observations with undue influence are excluded, using the DFBETA routine in *Stata*. Cluster-robust standard errors (in brackets) are clustered within primary studies. Wald Chi-square tests indicate overall significance of the hierarchical models (HM), which are also preferable to OLS given the log-likelihood for these models are smaller in magnitude.

In panel A, the PET/FAT results indicate substantial and positive selection bias as the *Constant* is larger than 1 and significant across three evidence pools. The presence of positive selection bias can be verified visually by inspecting the funnel graphs in Figure A1 in the *Appendix*. As indicated above, publication selection is a prevalent practice in medicine, physical sciences and social sciences. Results in Table 2 indicate that the research on R&D and productivity does not constitute an exception. It is beyond the scope of this study to discuss why publication selection bias exists despite the fact that the work of the leading contributors to this field was in great demand in the 1980s and 1990s. However, it is possible to conjecture that public policy makers are in need of justifying public support for private R&D investment on the basis of evidence demonstrating that the latter has direct or indirect positive effects on productivity of resident firms/industries. This public policy preference may be conducive to selection by researchers, who are interested in research uptake by and impact on the process of public policy-making. This conjecture draws support from evidence in the research field that both leading contributors and major reviews of the literature did tend to highlight representative/preferable estimates that indicate larger 'effects' compared to what the evidence indicates within each study and across studies.

However, the existence of selection bias does not invalidate the 'genuine' effect, which remains significant after controlling for selection bias. Therefore, we report PEESE results in Panel B of Table 2. The average estimate is 0.079 for elasticities in the level dimension; 0.057

for elasticities in the temporal dimension; and 11.7% for rates of return. These are still smaller than simple averages or representative/preferred estimates reported in previous reviews.[8]

Two points are worth emphasizing here. First, our findings confirm the consensus view that elasticity estimates in the temporal dimension are smaller and may be inconsistent (Hall et al., 2010; Hall and Mairesse, 1995). These estimates are likely to be influenced by collinearity between capital (both R&D and physical capital) and the time-effect that reflects autonomous technical change; and by measurement errors that are amplified when the data is first-differenced.

A more striking aspect of our findings is that they indicate a *gross* private rate of return on R&D that is *smaller* than the typical depreciation rate (usually, 15%) assumed in primary studies. This finding raises doubt as to whether the rate-of-return estimates reported in primary studies do indeed measure what they are supposed to measure.

With the exception of the debate on the difference between private and social returns on R&D, the primary-study authors and reviewers do not question whether the private rate-of-return estimates measure what they are supposed to measure. This is despite the fact that Griliches and Mairesse (1991a) drew attention to the limitations of the rate-of-return estimates and characterised them as only "distant reflections" of the true rate-of-return measures for two reasons. First, the estimates are contemporaneous partial effects of the R&D intensity on output or TFP growth. This is a naïve measure because R&D projects take several years to complete and the returns on completed R&D projects may not materialise until a few years after completion. Second, the estimates are obtained from *R&D intensity* in one period only - in contrast to elasticity estimates based on past R&D capital stock and current R&D flows.

[8] The upward bias is particularly acute in Wieser (2005), who report an average elasticity estimate of 0.179 in the temporal dimension. However, our finding for rate-of-return estimate is closer to the 13.8% reported in Moen and Thorsen (2013) after correcting for selection bias.

Therefore, Griliches and Mairesse (1991a: 338) conjecture that private rates-of-return estimates obtained from microeconometric models tend to be biased downward by an order of 50%. Taken in conjunction with this warning, our rate-of-return estimate of 11.7% indicates that the current specifications of the primal model may not be adequate for obtaining correct rates-of-return estimates. To obtain such estimates, it is necessary to model the lag structure of the R&D investment explicitly; and to estimate 'long-run' rates of return that take account of the lag structure for the relationship between R&D investment and its effects on output/TFP growth.

4.3 Multivariate meta-regression results

In what follows, we investigate how the moderating factors affect the estimates reported in primary studies. We measure the moderating factors with dummy variables that capture a specific feature of the research field *vis-a-vis* a reference category.[9] We use a hierarchical model specification justified by LR tests and follow a general-to-specific model estimation routine discussed in the methodology. Specific- and general-model estimates are presented side by side with a view to establish whether the findings are stable and congruent across models.

In the upper half of Table 3, we present coefficient estimates for a range of moderating variables that capture six dimensions of the research field: (i) publication type; (ii) measurement of output and inputs; (iii) model specification; (iv) estimation method; (v) country of origin for the data; and (vi) sample-related issues. At the bottom of Table 3, LR tests indicate that the hierarchical models fit the data better than their standard linear counterparts.

Before we discuss the findings for each dimension, we first discuss the range of *insignificant* moderating variables. In the **publication type dimension**, neither journal articles nor working papers are associated with systematically different elasticity estimates in the level

[9] Descriptions and summary statistics of the moderating variables are given in the Table A1 and Table A2 in the Appendix. Only one covariate (the method for constructing R&D capital) does not feature in all models as rates-of-return estimates are based on R&D intensity rather than R&D capital.

or temporal dimensions. However, there is evidence that journal articles tend to report smaller rate-of-return estimates. These findings suggest that journal articles does not suffer from the ‘winner’s curse’, which arises when journals with higher levels of perceived quality capitalise on their reputations and publish more selected results (Costa-Font et al., 2013). We also find that the country of origin for the data is usually insignificant, with the exception of US data which we discuss below. Lack of systematic difference between France, Germany, or the UK on the one hand and the rest of OECD as the reference category may be driven by similar levels of R&D intensity or by rate-of-return equalisation or both. With respect to the **sampling dimension**, we find that the median year in the time dimension of the panel data, the use of firm as opposed to industry data or the restriction of the sample to small firms as opposed to large or mixed-size firms do not explain the variation in elasticity or rate-of-return estimates.

As indicated above, US data is associated with relatively larger elasticity estimates in the level dimension but lower rate-of-return estimates; and small-firm data is associated with relatively larger elasticity estimates in the temporal dimension. Larger elasticity estimates associated with US Data is likely to be driven by higher R&D intensity in the US, which tops the OECD group throughout the time period covered in primary studies. This finding is congruent with country-level evidence reported in Soete and Verspagen (1993) and Coe and Helpman (1995) – and can be explained by higher technological capacity that enables R&D-intensive firms/industries to better exploit the productivity gains from R&D investment. On the other hand, lower rate-of-return estimates associated with US data is likely to be due to diminishing returns on R&D investment in countries where R&D intensity is higher than the reference category.

Table 3: Multivariate meta-regression results: sources of heterogeneity
(Dependent variable: t-values)

	Level elasticities		**Temporal elasticities**		**Rates of return**	
	General	Specific	General	Specific	General	Specific
Precision	0.043	0.044**	0.064	0.019	0.170***	0.167***
	(0.031)	(0.019)	(0.049)	(0.019)	(0.048)	(0.031)
Publication type						
Journal article	0.024		-0.092**		-0.117***	-0.116***
	(0.031)		(0.042)		(0.035)	(0.029)
Working paper	0.035		-0.066		0.040	
	(0.030)		(0.046)		(0.050)	
Measurement issues						
Output measured as value added	0.040***	0.045***	0.026***	0.025***	0.046**	0.038***
	(0.010)	(0.009)	(0.006)	(0.006)	(0.022)	(0.014)
Non-PIM method for R&D capital	-0.050**	-0.060***	-0.071*		NA	NA
	(0.023)	(0.021)	(0.042)		NA	NA
Control for double counting	0.019***	0.019***	0.029***	0.030***	-0.137***	-0.123***
	(0.003)	(0.003)	(0.007)	(0.007)	(0.049)	(0.034)
Weighted variables	-0.001		-0.027		0.102***	0.100***
	(0.009)		(0.075)		(0.029)	(0.019)
Model specification						
Enhanced panel-data models	0.050**	0.046***	-0.004		0.028	
	(0.022)	(0.015)	(0.020)		(0.059)	
Control for spillovers	0.012		0.030*		-0.035**	-0.027*
	(0.023)		(0.017)		(0.017)	(0.016)
Time dummies included	-0.009		0.037***	0.033***	0.021	0.042**
	(0.011)		(0.011)	(0.011)	(0.026)	(0.018)
Industry dummies included	-0.019***	-0.019***	0.063***	0.064***	-0.010	
	(0.004)	(0.004)	(0.004)	(0.004)	(0.029)	
Variable returns to scale allowed	-0.003		-0.042***	-0.044***	0.018	
	(0.003)		(0.009)	(0.009)	(0.025)	
Translog production function	-0.040		0.070		0.134**	0.133***
	(0.033)		(0.063)		(0.061)	(0.039)
Estimation issues						
Instrumented variable estimation	-0.059***	-0.054***	-0.034**	-0.043***	-0.006	
	(0.009)	(0.007)	(0.015)	(0.011)	(0.019)	
GMM estimation	-0.006		-0.037		-0.036	
	(0.023)		(0.025)		(0.067)	
Country of origin						
French firm or industry data	0.001		0.021		0.013	
	(0.010)		(0.030)		(0.049)	
German firm or industry data	-0.014		0.003		0.055	
	(0.027)		(0.048)		(0.061)	
UK firm or industry data	-0.025		0.061		0.034	
	(0.024)		(0.075)		(0.045)	
US firm or industry data	-0.019		0.057**	0.037***	-0.028	-0.042**
	(0.014)		(0.028)	(0.011)	(0.027)	(0.018)
Sampling issues						
Data mid-point after 1980	-0.004		-0.006		0.013	
	(0.008)		(0.014)		(0.016)	

Firm-level data	0.028	0.032*	0.017		-0.026	
	(0.020)	(0.019)	(0.032)		(0.026)	
R&D-intensive firms/industries	0.023***	0.024***	0.042***	0.043***	0.061	0.080*
	(0.009)	(0.009)	(0.011)	(0.011)	(0.052)	(0.048)
Small firms	-0.012		-0.064***	-0.065***	0.068	
	(0.007)		(0.018)	(0.018)	(0.050)	
Government-funded R&D	-0.104**	-0.099**	-0.143***	-0.145***	-0.196***	-0.195***
	(0.045)	(0.044)	(0.032)	(0.033)	(0.033)	(0.033)
Constant	1.431**	1.291**	0.932***	0.866***	1.309***	1.368***
	(0.574)	(0.562)	(0.299)	(0.302)	(0.208)	(0.187)
Random effects						
Standard dev. of random slopes	0.041***	0.043***	0.039***	0.049***		
Standard dev. of random intercepts	2.094**	2.056**	0.000***	0.000	0.633*	0.608**
Standard dev. of residuals	3.093***	3.124***	3.364***	3.360***	1.551***	1.573***
Model diagnostics						
Observations	440	440	468	468	350	350
Number of studies (clusters)	30	32	34	30	32	34
Model degrees of freedom	24	10	24	11	23	11
Log likelihood (hierarchical model)	-1175	-1180	-1174	-1180	-665	-669
Chi-square	355	334	424	409	214	201
P > Chi-square	0	0	0	0	0	0
Log likelihood (comparison model)	-1298	-1350	-1223	-1322	-672	-677

Notes: NA indicates not applicable. ***, ** and * indicate significance at 1%; 5% and 10%, respectively. Significance of the random-effect components are based on the natural log of the standard deviations. Results for elasticity estimates in the level and temporal dimension are based on a model with random intercepts and random slopes; those for rates-of-return estimation are based on a model with random intercepts only. Model choices are based on LR tests. Observations with undue influence are excluded, using the DFBETA routine in *Stata*. Models are not estimated with cluster-robust standard errors because the number of restrictions in the general model exceeds the number of clusters. However, cluster-robust estimation of the specific model is available on request. The results are the same – with the exception of firm-level data and R&D-intensive covariates, which become insignificant in the cluster-robust estimation of level dimension model. Wald Chi-square tests indicate overall significance of the hierarchical models that, according to the log-likelihood values, are also preferable to their standard linear counterparts.

Now we return to dimensions of the research field where a range of moderating variables tend to have a significant effect on the elasticity and rate-of-return estimates. Four dimensions stand out: **measurement of output and inputs**; **sample-related issues**; **model specification**; and **estimation methods**.

With respect to **input/output measurement**, we find that studies that measure output with value added tend to report larger elasticity and rate-of-return estimates. This is an interesting result given the lack of consensus in the literature on whether the use of value added affects reported estimates. Some studies report no difference between elasticity estimates based on value added and those based on sales corrected for intermediate inputs (Cunéo and Mairesse, 1984; Mairesse and Hall, 1994). Some others indicate that elasticity estimates based on value added tend to be smaller than those based on sales not corrected for intermediate inputs (Griliches and Mairesse, 1981).

Hall et al. (2010) indicate that the preferred measure is gross output, used in conjunction with intermediate inputs, capital and labour. However, they also cite two reasons as to why value added should be preferred to gross output or sales, particularly when the analysis is at the firm level. First, the ratio of materials to gross output can vary substantially across firms because of different degrees of vertical integration. Secondly, when output is measured by sales or gross output, the demand for intermediate inputs should be modelled explicitly, including the adjustment costs related to stocking of materials. These conditions are usually not satisfied due to data limitations.

We argue that larger elasticity estimates associated with the use of value added is a result to be expected. Note that the elasticity estimate is $\gamma = \rho(K_{it}/Y_{it})$, where ρ is the rate of return on R&D capital, and Y_{it} and K_{it} are sample means of output and R&D capital, respectively. When measured as value added, output is smaller than sales or gross output including cost of materials. Therefore, the value of (K_{it}/Y_{it}) is relatively larger and hence the elasticity (γ) is also relatively larger for the same rate of return (ρ). A similar explanation holds for larger rate-of-return estimates. Recall from equations (3b) and (3c) that the rate-of-return estimate (ρ) is the coefficient on R&D intensity (R_{it}/Y_{it}), which is larger when output is measured as value added rather than sales or gross output.

Schankerman (1981) is the first study that quantifies the downward bias in elasticity estimates when physical capital and labour inputs are not corrected for double counting. The bias is larger the larger are the ratios of R&D capital and R&D personnel to conventional capital and labour, respectively. Our finding confirms that, unless R&D investment is capitalised in national accounts, it is good practice to deduct the R&D capital and R&D labour from physical capital and total employment.[10] This is the case too when rates of return are estimated indirectly, using the elasticity estimate (γ). However, correcting for double counting introduces a downward bias in the rates of return estimated *directly*. This is to be expected because direct rates-of-return estimates are based on R&D intensity rather than R&D capital as an additional input.

Primary studies that use other methods to construct the R&D capital tend to report smaller elasticity estimates compared to those using the perpetual inventory method (PIM). Although this finding is limited to elasticities in the level dimension, it is worth highlighting here because the appropriate method is a contentious issue in the literature. PIM is compatible with the neo-classical theory of capital, which assumes that firms within a given industry will carry out less R&D investment in the current period if they have relatively higher levels of R&D capital stock in the preceding period. However, this assumption is usually not supported by empirical evidence – which indicates that firms that carry out above-average levels of R&D investment in the preceding period tend to do so in the current period too (Hall et al., 1986; Klette, 1994). Therefore, alternative methods for constructing R&D capital are recommended (Hall and Hayashi 1989; Klette, 1994; Bitzer and Stephan, 2007).

Although primary studies and Hall et al. (2010) discuss the merits and demerits of different methods for constructing the R&D capital, no systematic evaluation has been provided about whether productivity estimates would differ between studies using different methods. We

[10] Double counting will be less of an issue under the European System of Accounts (ESA) introduced in 2010. Under ESA (2010), R&D investment will no longer be treated as intermediate consumption. Instead, it will be treated as investment in intangible assets.

address this question here and report a downward bias in elasticity estimates when primary studies use other methods instead of PIM. This is due to the fact that the majority of the alternative methods are based on R&D capital proxies such as proportion of researchers to total employment or R&D investment per employee.[11] These measures do not correct for the empirical pattern that seems to be in contradiction with the assumption underlying the PIM– namely the positive correlation between R&D capital stocks in the last and current period. Therefore, we suggest that future research should limit innovation to more innovative alternatives as suggested by Klette (1994) rather than simple proxies for R&D capital.

Some studies use weighted variables when they estimate rates of return on R&D. The weight could be the square-root of the R&D intensity or firm size or industry's share in sectoral value added (Bartelsman, 1996; Cameron et al., 2005; Hall, 1993; and Lichtenberg and Siegel, 1991). Bartelsman et al. (1996) report that weighted estimations yield lower elasticity but higher rates-of-return estimates; but others do not provide comparative findings. Our finding corroborates Bartelsman et al. (1996) with respect to larger rate-of-return estimates; but not with respect to smaller elasticity estimates in the level or temporal dimensions. Therefore, we suggest that researchers and research users should compare rate-of-return estimates based on both weighted and un-weighted specifications.

With respect to **sampling issues**, we find that data on R&D-intensive firms/industries is associated with larger elasticity estimates in both dimensions and with higher rates of return. This is in line with several findings in primary studies (Griliches, 1980b; Griliches and Mairesse, 1981; Cunéo and Mairesse, 1984; Odagiri, 1983; Bartelsman, 1990; and Hall, 1993) and with the conclusions derived in the narrative synthesis by Hall et al. (2010). The standard explanation in the literature is that R&D-intensive firms/industries have better technological

[11] Other methods include the declining balance method with variable depreciation rates in Hall (1993); the rate of growth of R&D investment instead of R&D capital in Griliches (1980b); log of R&D investment instead of log of R&D capital in Bond et al (2002) and Rogers (2010); R&D expenditures per employee in Griffith et al (2006); and proportion of researchers to total employment in Ballot et al (2006).

capacity to exploit the benefits of the product and process innovations that R&D investment generates.

We also find smaller elasticity (and rate-of-return) estimates when the underlying data is government-funded R&D instead of private-funded or total R&D. Bartelsman (1990), Lichtenberg and Siegel (1991), Mansfield (1980), Terleckyj (1980), and Wolff and Nadiri (1993) all report smaller elasticity and rate-of-return estimates for government-funded R&D. Hall et al. (2010) suggest several reasons for the difference. First, firms may underestimate the risks when they use public funds for R&D purposes. Second, public funds for R&D may be spent in areas such as health and defence, with high levels of externalities. Finally, government funding of R&D may be concentrated in few industries (such as pharmaceuticals and IT) where returns are lower due to high levels of R&D intensity.

We are of the view that these sample-related findings point out two important issues that do not feature sufficiently in the existing literature and its reviews. The first concerns the role of market power in the case of larger productivity effects in R&D-intensive firms/industries. The question is: are larger elasticity and rate-of-return estimates in R&D-intensive firms/industries due to better technological capabilities or higher market power that enables them to extract innovation rents? The relationship between innovation and market power has been discussed extensively in the industrial organisation literature (see, Gilbert, 2006 for a review) but not in the R&D and productivity literature. The evidence from the former indicates that R&D intensity and market power are correlated positively – at least until a threshold of market power is reached (Aghion et al., 2005). Therefore, controlling for market power and/or for interactions between market power and R&D-intensity would constitute useful avenues for future research.

Secondly, and as indicated by Griliches (1979), decomposing the R&D capital into public and private components raises the issue of functional form in the primal model, where all inputs are assumed as complements and as such it is legitimate to include each input separately. However, is complementarity also applicable to the components of the R&D capital itself (e.g., private *versus* public R&D capital or basic *versus* applied R&D capital)? Or should these components be considered as substitutes, summed up into a single aggregate measure of R&D capital? These questions cannot be answered without testing for functional form, which is usually not the case in existing studies. Hence, further research is necessary to ascertain whether productivity differentials between public and private R&D are genuine or reflect a model specification bias.

With respect to **estimation methods,** the most contentious issue is how to address endogeneity and whether addressing endogeneity yields systematically different estimates. Some studies address endogeneity through a semi-reduced form of the production function (Griliches and Mairesse, 1981). Some others use three-stage least-squares (3SLS) (Verspagen, 1995) or a general method of moments (GMM) estimator (Mairesse and Hall, 1996; Aldieri et al., 2008; Blanchard et al., 2006; and Griffth et al., 2006). Yet, there is no consensus about how the instrumental variable estimation methods would affect the reported estimates (see, Hall et al., 2010). Our findings indicate that studies that implement an instrumental variable estimation method (2SLS, 3SLS, or system/difference GMMM) tend to report smaller elasticity estimates in both level and temporal dimensions. The effect is also negative but insignificant in the rate-of-return estimates. Therefore, the assumption that different estimators produce statistically similar results is at best questionable (Eberhardt and Helmers, 2010). However, it must also be noted that results from instrumental variable estimators (including GMM) are highly contingent on specifying the correct instrument mix.

The largest number of moderating factors with significant effects is observed within the **model specification** dimension. However, the findings here are partial in the sense that they relate to a single evidence base (elasticities in the level or temporal dimensions or rate-of-return estimates only). Following the order of reporting in Table 3, we observe that studies that use Enhanced panel-data models that control for panel co-integration and cross-sectional dependence report larger elasticity estimates in the level dimension. Notable among these studies are Anon and Higon (2007) who utilise an autoregressive distributed lag (ARDL) model; Doraszelski and Jaumandreu (2013) who utilise a controlled Markov process that captures the impact of R&D on the evolution of productivity; and Eberhardt et al (2013) who take account of cross-sectional dependence by using common correlated effects pooled estimator (CCEP) or common correlated effects mean group estimator (CCEMG). These are innovative approaches but the larger estimates they report are significant only in the level dimension.

A large body of empirical work investigates whether R&D spillovers has *indirect* productivity effects. However, there is no systematic evaluation of whether the *direct* effects differ between studies that do or do not control for the spillover effects separately. Our finding indicates no systematic difference when studies estimate elasticities, but rate-of-return studies that control for spillovers tend to report smaller rates-of-return estimates compared to others that do not.

Unlike existing reviews, however, we argue that this finding should be taken with a pinch of salt because the way in which spillovers are measured and modelled begs important questions – as indicated emphatically in a working paper by Griliches (1991). First, primary studies tend to report contemporaneous spillover effects despite the fact that spillovers are likely to take more time than own R&D to have an effect on productivity. Secondly, it is difficult to measure the pool of external R&D capital with precision either through weighted measures where the weights usually consist of technology proximity matrix or unweighted measures where the spillover effects is assumed to be symmetric across industries. Finally, the primary studies tend to assume that all firms would benefit from the spillover pool equally, whereas heterogeneity is more likely due to cross-firm differences in R&D intensity as a potential determinant of absorption capacity.

The third partial finding indicates that primary studies that include industry dummies in their models tend to report smaller elasticity estimates in the level dimension. This is in line Hall et al. (2010), who report that elasticities in the level dimension tend to be smaller when primary studies include industry/sector dummies in their models. In the temporal dimension, however, we find a positive coefficient. With respect to time dummies, we find that their inclusion in primary-study models is associated with larger elasticity estimates in the temporal dimension and with higher rate-of-return estimates.

Inclusion of industry/time dummies is considered as good practice that takes account of unobservable factors such as quality differences not reflected in prices (Hall et al., 2010). However, there is little or no guidance on correcting for new sources of bias that the inclusion of time/industry dummies may introduce when the latter are correlated with existing covariates. Therefore, we recommend that researchers should complement the use of time/industry dummies with estimations based on industry subsets characterised by technological proximity. Pavitt's (1984) taxonomy of technology classes can be a useful framework for such analysis.

The fourth set of partial findings indicate that primary studies that allow for variable returns to scale tend to report smaller elasticity estimates in the temporal dimension. This is in line with Hall et al. (2010), who report that studies imposing constant returns to scale tend to report larger elasticity estimates in the temporal dimension. It is also in line with a number of primary studies that report a similar pattern, including Griliches and Mairesse (1981), Cunéo and Mairesse (1984), and Griliches and Mairesse (1991a). However, when we checked the primary studies allowing for variable returns to scale we find out that the lower elasticity estimates they report are associated with decreasing return to scale. Hence, the lower elasticity estimates in studies allowing for variable returns to scale are likely to be driven by decreasing returns to scale. However, Cunéo and Mairesse (1984: 378) indicate that the decreasing returns to scale they establish are somewhat implausible. Therefore, the reliability of the elasticity estimates in the temporal dimension is questionable not only because of the amplification of the measurement error discussed above, but also because of their susceptibility to whether the data at hand indicates increasing or decreasing returns to scale.

Conclusions

The work on R&D and productivity has made significant contributions to knowledge not only in terms of empirically-rich findings but also with respect to measurement, modeling and estimation issues involved. Our review has enabled us to take stock of the findings from the extant literature and provide verifiable findings on productivity effects of R&D investment, including elasticity estimates in the level and temporal dimensions and rates-of-return estimates. Some of our findings are in line with those reported in a number of narrative reviews and meta-analysis studies. Congruent findings include: (i) the average productivity effect of R&D investments is positive – whether it is measured as output elasticity or rate of return; (ii) controlling for double-counting of the R&D capital and R&D personnel is necessary to avoid a downward bias in the elasticity estimates; (iii) elasticity estimates in R&D-intensive firms/industries are usually larger than other industries, and (iv) elasticity and rate-of-return estimates for government-funded R&D are lower compared to private-funded R&D.

However, even with respect to congruent findings, we provide additional evidence indicating that the findings in existing reviews should be qualified. Specifically: (i) the productivity effect may be positive, but it is smaller than what is reported in existing reviews and the ‘average’ rate-of-return estimate is smaller than the rate of depreciation usually assumed; (ii) controlling for double counting may be appropriate in the estimation of elasticities, but it is likely to introduce a downward bias in rates-of-return estimates; (iii) the productivity effects of R&D investment may be larger in R&D-intensive firms/industries, but the cause can be either better technological capabilities among R&D-intensive firms/industries as suggested by existing reviews or higher market power as indicated in the industrial organisation literature; and (iv) the productivity effects of public R&D may be smaller than privately-funded R&D, but this results raises the question of functional form and is contingent on whether private and public R&D are complements rather than substitutes.

Beyond these qualifications, we also provide verifiable evidence about the effects of a range of moderating factors with respect to which the existing reviews provide either inconclusive or divergent conclusions. Here, we find that: (i) elasticity and rate-of-return estimates based on value added as the measure of output are larger than those based on gross output or sales; (ii) elasticity estimates that do not take account of endogeneity in the level or temporal dimensions are likely to be biased upward; (iii) elasticity estimates based on other methods of constructing the R&D capital stock are smaller than those based on the perpetual inventory method (PIM) and are likely to suffer from downward bias; (iv) elasticity estimates based on small-firm data are smaller than those for large firms in the temporal dimension, but this difference is likely to reflect a downward bias due to measurement error in small-firm data, which is exacerbated in the temporal dimension; (v) studies that control for spillovers as an additional source of productivity report smaller rates-of-return estimates compared to others that do not, but the existing methods through which spillovers are measured and modeled should be questioned; and (vi) variable returns to scale are associated with smaller elasticity estimates in the temporal dimension, but this association is likely to be due to decreasing returns to scale rather than relaxing the assumption of constant returns to scale.

Our findings indicate that future research could benefit from a number of innovations in modeling and estimating the R&D-productivity relationship. One innovation relates to rate-of-return estimations, where it is necessary to account for the lag structure in the R&D-productivity relationship and for the time-lag involved in completion of R&D projects. Secondly, relatively larger productivity effects associated with R&D-intensive firms poses the question as to whether this association is due to variations in absorption capacities or market power or both. Therefore, we recommend that the primal model be augmented with covariates capturing technology absorption capacity and market power/concentration measures, preferably with interaction terms between the latter and R&D intensity. Third, relatively smaller elasticity and rate-of-return estimates associated with government-funded R&D may be due to concentration of public support in R&D-intensive industries with high levels of externality; but they may also reflect a model specification bias due to untested assumption of complementarity between public and private R&D. Therefore, it is necessary to test for complementarity before reporting separet estimates for the productivity effects of private and public R&D. Finally, we agree with Hall et al. (2010) that there is scope for innovation in the methods used for constructing R&D capital; however the current practice is likely to constitute an additional source of bias as it consists of using simple proxies rather than alternative methods that challenge the neo-classical assumption of within-industry equalisation of R&D capital.

References

1. Work cited in the text (Except studies included in the meta-analysis)

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt. (2005). Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics*, 120(5), 701-728.

Bond, S., and I. Guceri (2012). Trends in UK BERD after the Introduction of R&D Tax Credits. Oxford University Centre for Business Taxation, Working Paper No. 12/01.

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

Coe, D. T. and E. Helpman (1995). International R&D spillovers. *European Economic Review*, *39*(5), 859-887.

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

Crépon, B., E. Duguet, and J. Mairesse (1998). Research, innovation, and productivity: An econometric analysis at the firm level, *Economics of Innovation and New Technology* 7, 115-156.

Cunéo, P. and J. Mairesse (1984). Productivity and R&D at the firm level in French manufacturing. In: Z. Griliches (ed.), *R&D, Patents and Productivity*. Chicago, IL: University of Chicago Press, 375- 392.

Demidenko, E. (2004). *Mixed Models: Theory and Applications*. Hoboken, NJ: Wiley.

Dickersin, K., & MIN, Y. I. (1993). Publication bias: the problem that won't go away. *Annals of the New York Academy of Sciences*, *703*(1), 135-148.

Doucouliagos, C(H) and T. D. Stanley (2009). Publication selection bias in minimum-wage research? A meta-regression analysis, *British Journal of Industrial Relations*, **47(2)**, 406-429.

Doucouliagos, H(C) and T. D. Stanley (2013). Theory competition and selectivity: Are all economic facts greatly exaggerated?, *Journal of Economic Surveys*, 2013; **27**: 316-39.

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

Eberhardt, M. and C. Helmers (2010). Untested assumptions and data slicing: A critical review of firm-level production function estimators. Department of Economics Discussion Paper no. 513, University of Oxford.

Egger, M., G. D Smith, M. Scheider, and C. Minder (1997). Bias in meta-analysis detected by a simple, graphical test, *British Medical Journal*, **316**: 629-34.

Gilbert, R. (2006), 'Looking for Mr. Schumpeter: Where are we in the competition-innovation debate?', in Adam B. Jaffe, Josh Lerner and Scott Stern (eds), *Innovation Policy and the Economy – Volume 6*, Cambridge, Mass.: MIT Press, pp. 159-215.

Goto, A. and K. Suzuki (1989). R&D capital, rate of return on R&D investment and spillover of R&D in Japanese manufacturing industries, *Review of Economics and Statistics*, 71(4), 555-564.

Griliches, Z. (1973). Research expenditures and growth accounting. In: *Science and Technology in Economic Growth*, R.B. Williams (ed.). John Wiley and Sons, New York.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth, *Bell Journal of Economics*, 10(1), 92-116.

Griliches, Z. (1980a). Returns to research and development expenditures in the private sector. In: *New Developments in Productivity Measurement and Analysis*, Ed. By J. W. Kendrick, and B. N. Vaccara. Chicago, Ill.: Chicago University Press, 419-462.

Griliches, Z. (1991). The search for R&D spillovers. National Bureau of Economic Research Working Paper, no. w3768.

Griliches, Z. (1994). Productivity, R&D and the data constraint, *American Economic Review*, 84(1), 1-23.

Griliches, Z. (1995). R&D and productivity: Econometric results and measurement issues, in P. Stoneman (ed.), *Handbook of the Economics of Innovation and Technical Change*, Blackwell Handbooks in Economics.

Griliches, Z. (1998). *R&D and productivity: The econometric evidence*. Chicago, IL: University of Chicago Press.

Griliches, Z., and F. R. Lichtenberg (1984). R&D and productivity growth at the industry level: Is there still a relationship? In: *R&D, Patents, and Productivity*, Ed. By Z. Griliches. Chicago, Ill.: Chicago University Press, 465-501.

Hall, B. H. (1993). Industrial research during the 1980s: Did the rate of return fall?, *Brookings Papers On Economic Activity*, Micro (2), 289-344.

Hall, B. H. (1996). The private and social returns to research and development. In Bruce L.R. Smith and Claude E. Barfield (eds.), *Technology, R&D, and the Economy*, Brookings Institution, Washington D.C.

Hall, B. H., J. Mairesse and P. Mohnen (2010). Measuring the returns to R&D. In B. H. Hall and N. Rosenberg (eds.), *Handbook of the Economics of Innovation*, vol. 2, Elsevier, New York.

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

Ioannidis, J. P. (2005). Contradicted and initially stronger effects in highly cited clinical research. *Jama*, *294*(2), 218-228.

Klette, T. J. (1994). R&D, scope economies, and company structure: a 'not so fixed effect' model of plant performance. Oslo, Norway: Central Bureau of Statistics *Discussion Paper* No. 120.

Klette, T. J. (1996). The accumulation of R&D-Capital and the dynamic performance of manufacturing firms. Oslo, Norway: Central Bureau of Statistics.

Mairesse, J. and M. Sassenou (1991). R&D and productivity: a survey of econometric studies at the firm level, *STI Review*, OECD, 8, 9-46.

Mairesse, J., and P. Mohnen (1994). R&D and productivity growth: What have we learned from econometric studies?, *Eunetic Conference on Evolutionary Economics of Technological Change: Assessment of Results and New Frontiers*, 817-888, Strasbourg.

Mansfield, E. (1980). Basic Research and Productivity Increase in Manufacturing, *American Economic Review*, 70, 863-873.

McCulloch, C. E., S. R. Searle, and J. M. Neuhaus (2008). *Generalized, Linear, and Mixed Models (2nd ed)*. Hoboken, NJ: Wiley.

Minasian, Jora R (1969). Research and development, production functions, and rates of return. *American Economic Review,* 59(2): 80-85.

Møen, J., & Thorsen, H. S. (2013). Publication bias in the returns to R&D literature, Institutt for Foretaksøkonomi *Discussion Paper*, No. 2013/12.

Moreno, S.G., A. J. Sutton, A. Ades, T. D. Stanley, K. R. Abrams, J. L. Peters, and N. J. Cooper (2009). Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study, *BMC Medical Research Methodology*, 9(2), 1–17.

OECD (2011). Business R&D. In OECD Science, Technology and Industry Scoreboard, OECD Publishing. http://dx.doi.org/10.1787/sti_scoreboard-2011-18-en

Pakes, A., and M. Schankerman (1984). The rate of obsolescence of patents, research gestation lags, and the private rate of return to research resources. in Z. Griliches (ed.), *R&D, Patents, and Productivity*, Chicago, IL: Chicago University Press, 73-88.

Park, W. G. (1995). International R&D spillovers and OECD economic growth, *Economic Inquiry* 33, 571-591.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, *22*(11), 1359-1366.

Soete, L. and B. Verspagen (1993). Convergence and divergence in growth and technical change: an empirical investigation. Paper presented at the AEA annual meeting in Anaheim, Ca., January.

Stanley, T. D. (2005). Beyond publication bias, *Journal of Economic Surveys,* **19(3),** 309-45.

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

Stanley, T.D. and H(C) Doucouliagos (2014). Meta-regression approximations to reduce publication selection bias, *Research Synthesis Methods*, *5***(1)**: 60-78.

Stanley, T.D. and H(C). Doucouliagos (2007). Identifying and correcting publication selection bias in the efficiency-wage literature: Heckman meta-regression, Deakin University Economics *Working Paper* No. 2007_11; 2007.

Stanley, T.D. and H. Doucouliagos (2012). *Meta-regression analysis in Economics and Business*. London and New York: Routledge.

Stanley, T.D. and S. B. Jarrell (1989). Meta-regression analysis: A quantitative method of literature surveys, *Journal of Economic Surveys,* **3(2)**: 161-170.

Terleckyj, N. E. (1974). *Effects of R&D on the productivity growth of industries: An exploratory study* (No. 140). Washington, DC: National Planning Association.

2. *Studies included in the meta-analysis*

1. Aiello, F. and P. Cardamone (2005). R&D spillovers and productivity growth: Evidence from Italian manufacturing microdata, *Applied Economics Letters*, 12(10), 625-631.
2. Aldieri L, M. Cincera, A. Garofalo, and C. P. Vinci (2008). Micro evidence of the effects of R&D on labour productivity for large international R&D firms, *International Journal of Manpower*, 29(3), 198-215.
3. Anon Higon, D. (2007). The impact of R&D spillovers on UK manufacturing TFP: A dynamic panel approach, *Research Policy*, 36(7), 964-979.
4. Ballot, G., F. Fakhfakh and E. Taymaz (2006). Who benefits from training and R&D: The firm or the workers?, *British Journal of Industrial Relations*, 44(3), 473-495.
5. Bartelsman, E. J. (1990). R&D spending and manufacturing productivity: An empirical analysis. Washington, DC*: Federal Reserve Board of Governors Finance and Economics Discussion* Series #122.
6. Bartelsman E. J., G. van Leeuwen, and H. Nieuwenhuijsen (1996). Advanced manufacturing technology and firm performance, *Netherlands Official Statistics* 11, 40-51.
7. Blanchard, P., J-P Huiban and P. Sevestre (2005). R&D and productivity in corporate groups: An empirical investigation using a panel of French firms, *Annals of Economics and Statistics*/*Annales d'Économie et de Statistique*, No. 79/80, 461-485.
8. Boler, E. A., A. Moxnes and K-H Ulltveit-Moe (2012). Technological change, trade in intermediates and the joint impact on productivity, *C.E.P.R. Discussion Papers.* No. 8884.
9. Bond S., D. Harhoff and J. van Reenen (2002). Corporate R&D and productivity in Germany and the United Kingdom, *Centre for Economic Performance Discussion Paper*s, LSE. No. 595.
10. Bonte, W. (2003). R&D and productivity: Internal vs. external R&D - Evidence from West German manufacturing industries, *Economics of Innovation & New Technology*. 12(4): 343-360.
11. Branstetter, L. (1996). Are knowledge spillovers international or intra-national in scope? Microeconometric evidence from the U.S. and Japan, NBER *Working Paper Series*. No. 5800.
12. Cameron, G., J. Proudman and S. Redding (2005). Technological convergence, R&D, trade and productivity growth, *European Economic Review*, 49(3), 775-807.
13. Cincera, M. (1998). *Economic and Technological Performances of International Firms.* PhD Thesis Presented to the Faculty of Social, Political and Economic Sciences, Free University of Brussels, 1998.
14. Clark, K. and Z. Griliches (1998). Productivity growth and R&D at the business level: Results from the PIMS data base. In: *R&D and Productivity: The Econometric Evidence*. University of Chicago Press, 134 - 156.
15. Cuneo, P. and J. Mairesse (1984). Productivity and R&D at the firm level in French manufacturing. In: Griliches Zvi (ed.), *R&D, Patents, and Productivity*. University of Chicago Press, 375 - 392.

16. Doraszelski, U. and J. Jaumandreu (2013). R&D and productivity: Estimating endogenous productivity, *The Review of Economic Studies*, 80(3): 1338-1383.
17. Eberhardt, M., C. Helmers and H. Strauss (2010). Do spillovers matter in estimation of private returns to R&D?, *Economic and Financial Reports*, European Investment Bank. No. 2010/01.
18. Frantzen, D. (2002). Intersectoral and international R&D knowledge spillovers and total factor productivity, *Scottish Journal of Political Economy*, 49(3), 280-303.
19. Goto, A. and K. Suzuki (1989). R &D capital, rate of return on R&D investment and spillover of R&D in Japanese manufacturing industries, *The Review of Economics and Statistics*, 71(4), 555-564.
20. Griffith, R., R. Harrison, and J. van Reenen (2006). How special is the special relationship? Using the impact of U.S. R&D spillovers on U.K. firms as a test of technology sourcing, *American Economic Review*, 96(5), 1859-1875.
21. Griffith, R., S. Redding, and J. van Reenen (2004). Mapping the two faces of R&D: Productivity growth in a panel of OECD manufacturing industries, *Review of Economics and Statistics,* 86(4), 883-895.
22. Griliches, Z. (1980a). Returns to research and development expenditures in the private sector. In John W Kendrick; Beatrice N Vaccara (eds), *New Developments in Productivity Measurement*. University of Chicago Press, 419-462.
23. Griliches, Z. (1980b) R&D and the Productivity Slowdown. *American Economic Review*. 70(2): 343-348.
24. Griliches, Z. (1998). Productivity, R&D, and basic research at the firm level in the 1970s. In Griliches Zvi, *R&D and Productivity: The Econometric Evidence*. University of Chicago Press, 82 - 99.
25. Griliches, Z., & Mairesse, J. (1981). Productivity and RD at the firm level. NBER Working Paper no. 826.
26. Griliches, Z. and J. Mairesse (1991a). R&D and productivity growth: Comparing Japanese and U.S. manufacturing firms. In Hulten Charles R (ed.), *Productivity Growth in Japan and the United States*. University of Chicago Press, 317 - 348.
27. Griliches Z. and J. Mairesse (1991b). Comparing productivity growth: An exploration of French and U.S. industrial and firm data. In MacNil de Georges and Gordon Robert J (eds), *International Volatility and Economic Growth: The First Ten Years of The International Seminar on Macroeconomics*. Elsevier Science Publishers B.V., 45 - 82.
28. Griliches, Z. and F. R. Lichtenberg (1984a). Interindustry technology flows and productivity growth: A re-examination, *Review of Economics & Statistics* 66, 325-329.
29. Griliches, Z., and J. Mairesse (1981). Productivity and R&D at the firm level, NBER *Working Papers*, no. 826.
30. Hall, B. H. and J. Mairesse (1995). Exploring the relationship between R&D and productivity in French manufacturing firms, *Journal of Econometrics* 65, 263-293.
31. Hall, B. H., E. Mansfield and A. B. Jaffe (1993). Industrial research during the 1980s: Did the rate of return fall?, *Brookings Papers on Economic Activity*. Microeconomics. 1993(2), 289-344.

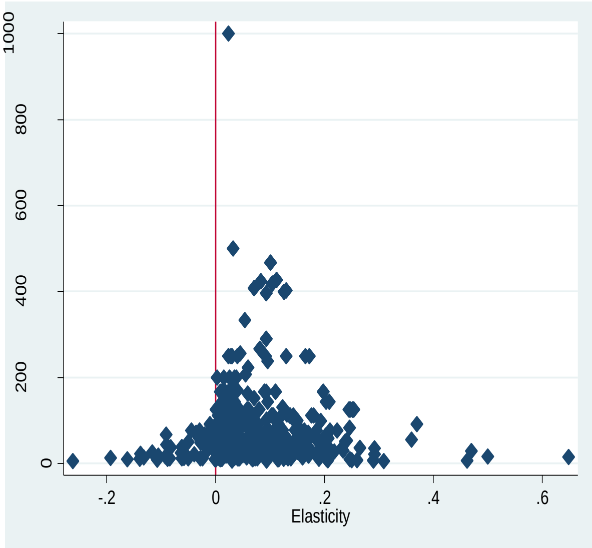
32. Hanel, P. (2000). R&D, interindustry and international technology spillovers and the total factor productivity growth of manufacturing industries in Canada, 1974–1989, *Economic Systems Research*, 12(3), 345-361.
33. Harhoff, D. (1994). R&D and productivity in German manufacturing firms. ZEW *Discussion Papers*, No. 94 – 01.
34. Harhoff, D. (2000). R&D Spillovers, Technological Proximity, and Productivity Growth - Evidence From German Manufacturing Firms, *Schmalenbach Business Review*, 52(3): 238-260.
35. Heshmati, A. and K. Hyesung (2011). The R&D and productivity relationship of Korean listed firms, *Journal of Productivity Analysis*, 36(2), 125-142.
36. Hsing, Y. and W-J Lin (1998). R&D spending and employee productivity in the United States, *International Journal of Management* 15: 79.
37. Kafouros, M. I. (2005). R&D and productivity growth: Evidence from the UK, *Economics of Innovation & New Technology*, 14(6), 479-497.
38. Klette, T. J. (1991). On the importance of R&D and ownership for productivity growth. Evidence from Norwegian micro-data 1976-85, *Central Bureau of Statistics Discussion Paper*. No. 60.
39. Kwon, H. U. and T. Inui (2003). R&D and productivity growth in Japanese manufacturing firms, *Economic and Social Research Institute* (*ESRI) Discussion Paper Series*. No. 44.
40. Lehto, E. (2007). Regional impact of Research and Development on productivity, *Regional Studies*, 41(5), 623-638.
41. Lichtenberg, F. R. and D. Siegel (1991). The impact of R&D investment on productivity: New evidence using linked R&D-LRD data, *Economic Inquiry*, 29(2): 203.
42. Link, A. N. (1981). Basic research and productivity increase in manufacturing: Additional evidence, *American Economic Review*, 71(5), 1111-1112.
43. Link, A. N. (1983). Inter-firm technology flows and productivity growth, *Economics Letters*, 11(1), 179-184.
44. Lokshin, B., R. Belderbos and M. Carree (2008). The productivity effects of internal and external R&D: Evidence from a dynamic panel data model, *Oxford Bulletin of Economics & Statistics*, 70(3), 399-413.
45. Los, B. and B. Verspagen (2000). R&D spillovers and productivity: Evidence from U.S. manufacturing microdata, *Empirical Economics*, 25(1), 127-148.
46. Mairesse J., and B. H. Hall (1996). Estimating the productivity of research and development: An exploration of GMM methods using data on French and United States manufacturing firms. NBER *Working Paper Series*. No. 5501.
47. Mate-Garcia J. J. and J. M. Rodriguez-Fernandez (2008). Productivity and R&D: An econometric evidence from Spanish firm-level data, *Applied Economics*, 40(14), 1827-1837.
48. Medda, G., C. Piga and D. Siegel (2003). On the relationship between R&D and productivity: A treatment effect analysis. : Fondazione Eni Enrico Mattei, *Nota Di Lavoro* 34.

49. O'Mahony, M. and M. Vecchi (2009). R&D, knowledge spillovers and company productivity performance, *Research Policy,* 38(1), 35-44.
50. Odagiri, H. (1983). R & D expenditures, royalty payments, and sales growth in Japanese manufacturing corporations. *The Journal of Industrial Economics*, 32(1), 61-71.
51. Odagiri H, and H. Iwata (1986). The impact of Research and Development on productivity increase in Japanese manufacturing companies, *Research Policy*, 15(1), 13-19.
52. Ortega-Argiles, R., M. Piva, L. Potters and M. Vivarelli (2010). Is corporate R&D investment in high-tech sectors more effective?, *Contemporary Economic Policy*, 28(3), 353–365.
53. Rogers, M. (2010). R&D and productivity: Using UK firm-level data to inform policy, *Empirica*, 37(3), 329-359.
54. Schankerman, M. (1981). The effects of double-counting and expensing on the measured returns to R&D, *Review of Economics and Statistics,* 63(3), 454-458.
55. Scherer, F. M. (1982). Interindustry technology flows and productivity growth, *Review of Economics and Statistics,* 64(4), 627-634.
56. Smith, V., M. Dilling-Hansen, T. Eriksson, E. S. Madsen (2004). R&D and productivity in Danish firms: Some empirical evidence, *Applied Economics*, 36(16), 1797-1806.
57. Sorensen, A., H. C. Kongsted and M. Marcusson (2003). R&D, public innovation policy, and productivity: The case of Danish manufacturing, *Economics of Innovation and New Technology*, 12(2), 163-78.
58. Sterlacchini, A. (1989). R&D, innovations, and total factor productivity growth in British manufacturing, *Applied Economics*, 21(11): 1549.
59. Sveikauskas, L. (1981). Technology inputs and multifactor productivity growth, *Review of Economics and Statistics*, 63(2), 275-282.
60. Terleckyj, N. (1980). Direct and indirect effects of industrial research and development on the productivity growth of industries. In J. Kendrick and B. Vaccara (eds.), *New Developments in Productivity Measurement and Analysis*. UMI/NBER publication, 357-386.
61. van Meijl, H. (1997). Measuring intersectoral spillovers: French evidence, *Economic Systems Research*, 9(1), 25-46.
62. Verspagen, B. (1995). R&D and productivity: A broad cross-section cross-country look, *Journal of Productivity Analysis*, 6(2), 117-35.
63. Verspagen, B. (1997). Estimating international technology spillovers using technology flow matrices, *Weltwirtschaftliches Archiv*, *133*(2), 226-248.
64. Wakelin, K. (2001). Productivity growth and R&D expenditure in UK manufacturing firms. *Research Policy*, 30(7), 1079-1090.
65. Wolff, E. N., and M. I. Nadiri (1993). Spillover effects, linkage structure, and Research and Development, *Structural Change and Economic Dynamics*, 4(2), 315-331.

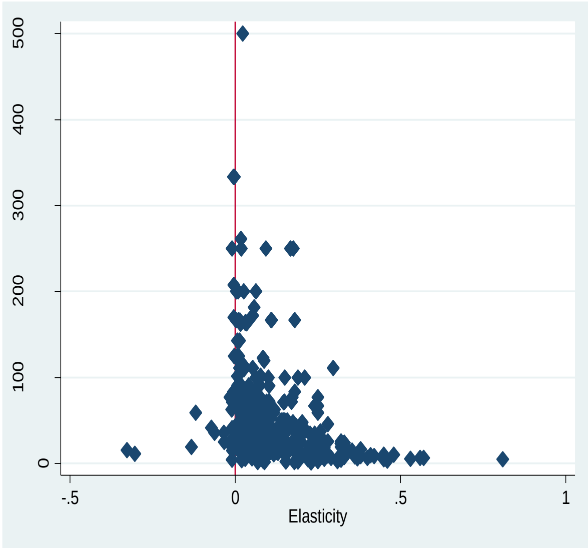
APPENDIX

Figure A1: Funnel graphs for visual detection of publication selection bias

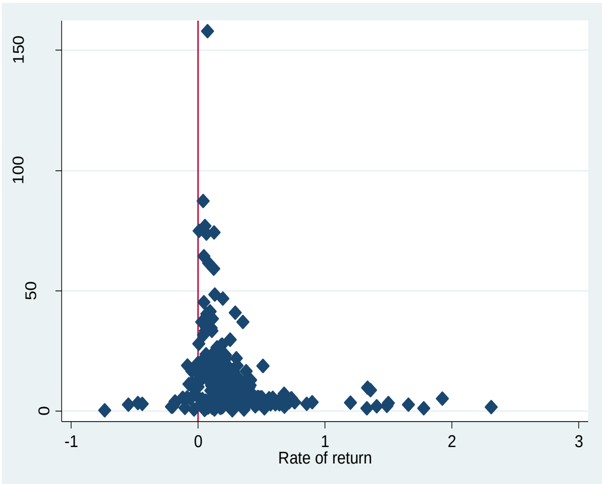
Elasticity estimates in the level dimension



Elasticity estimates in temporal dimension



Rate-of-return estimates



Note: the funnel graphs are generated after excluding observations with undue influence (observations with |DFBETA|>1)

Table A1: Moderating variables: descriptions and reference categories

Moderating variable	Description and reference category
1. Journal article	Equals 1 if study is published as journal article; 0 for all other publications types
2. Working paper	Equals 1 if study is published as working paper; 0 for all other publications types
3. Firm-level data	Equals 1 if estimate is based on firm- or plant-level data; 0 if it is based on 2-digit or more disaggregated industry data
4. GMM estimation	Equals 1 if estimate is based on system or difference GMM estimation; 0 for all other types of estimation
5. Weighted variables	Equals 1 if estimate is based on weighted variables; 0 otherwise
6. Enhanced panel-data models	Equals 1 if estimate is based on enhanced panel-data models such as panel cointegration and cross-sectional dependence; 0 otherwise
7. Instrumented variable estimation	Equals 1 if estimate is based on instrumented (GMM, 2SLS, 3SLS, etc.) estimation; 0 otherwise
8. Output measured as value added	Equals 1 if estimate is based on value added; 0 for output measured as sales or production
9. Data mid-point after 1980	Equals 1 if estimate is based on data panel with amid year = 1980; 0 if the midpoint is larger than 1980
10. French firm or industry data	Equals 1 if estimate is based on French data; 0 for data from ALL other OECD countries
11. German firm or industry data	Equals 1 if estimate is based on German data; 0 for data from ALL other OECD countries
12. UK firm or industry data	Equals 1 if estimate is based on UK data; 0 for data from ALL other OECD countries
13. US firm or industry data	Equals 1 if estimate is based on US data; 0 for data from ALL other OECD countries
14. Control for double counting	Equals 1 if the estimate is derived from a model that controls for double counting; 0 otherwise
15. Control for spillovers	Equals 1 if the estimate is derived from a model that controls for spillovers; 0 otherwise
16. Time dummies included	Equals 1 if the estimate is derived from a model that includes period dummies; 0 otherwise
17. Industry dummies included	Equals 1 if the estimate is derived from a model that includes industry dummies; 0 otherwise
18. Variable returns to scale allowed	Equals 1 if estimate is derived from a model that allows for variable returns to scale; 0 if constant returns are imposed
19. R&D-intensive firms	Equals 1 if estimate relates to R&D-intensive firms/industries as defined by the author; 0 otherwise

20. Small firms	Equals 1 if estimate relates to small firms as defined by the author; 0 otherwise
21. Translog production function	Equals 1 if estimate is based on translog production function; 0 otherwise
22. Government-funded R&D	Equals 1 if estimate is based on government-funded R&D; 0 otherwise
23. Non-PIM method for R&D capital	Equals 1 if estimate is NOT based on perpetual inventory method; 0 otherwise
24. Weighted variables	
	Equals 1 if estimate is related to level of government-funded R&D; 0 for private funded or source not specified

Table A2: Summary statistics for moderating variables

	Elasticities (level dimension)					Elasticities (temporal dimension)					Rates of return				
	Obs.	Mean	Std. Dev.	Min.	Max.	Obs.	Mean	Std. Dev.	Min.	Max.	Obs.	Mean	Std. Dev.	Min.	Max.
Effect size															
Elasticity / Return estimate	440	0.078	0.095	-0.262	0.648	468	0.109	0.114	-0.328	0.81	350	0.250	0.312	-0.737	2.310
Standard error	440	0.030	0.032	0.001	0.210	468	0.052	0.085	0.002	0.92	350	0.183	0.299	0.006	2.989
Publication type															
Journal article	440	0.661	0.474	0	1	468	0.387	0.488	0	1	350	0.780	0.415	0	1
Working paper	440	0.236	0.425	0	1	468	0.389	0.488	0	1	350	0.086	0.280	0	1
Measurement issues															
Output measured as value added	440	0.477	0.500	0	1	468	0.353	0.478	0	1	350	0.391	0.489	0	1
Non-PIM method for R&D capital	440	0.400	0.490	0	1	468	0.128	0.335	0	1	NA	NA	NA	NA	NA
Control for double counting	440	0.252	0.435	0	1	468	0.318	0.466	0	1	350	0.234	0.424	0	1
Weighted variables	440	0.080	0.271	0	1	468	0.011	0.103	0	1	350	0.131	0.338	0	1
Model specification															
Enhanced panel-data models	440	0.105	0.306	0	1	468	0.083	0.277	0	1	350	0.060	0.238	0	1
Control for spillovers	440	0.155	0.362	0	1	468	0.141	0.348	0	1	350	0.203	0.403	0	1
Time dummies included	440	0.543	0.499	0	1	468	0.504	0.501	0	1	350	0.320	0.467	0	1
Industry dummies included	440	0.457	0.499	0	1	468	0.192	0.395	0	1	350	0.463	0.499	0	1
Variable returns allowed	440	0.482	0.500	0	1	468	0.412	0.493	0	1	350	0.326	0.469	0	1
Translog production function	440	0.182	0.386	0	1	468	0.049	0.216	0	1	350	0.106	0.308	0	1
Estimation issues															

Instrum'd.variable estimation	440	0.268	0.444	0	1	468	0.115	0.320	0	1	350	0.183	0.387	0	1
GMM estimation	440	0.114	0.318	0	1	468	0.056	0.229	0	1	350	0.020	0.140	0	1
Country															
French firm or industry data	440	0.095	0.294	0	1	468	0.075	0.263	0	1	350	0.063	0.243	0	1
German firm or industry data	440	0.055	0.227	0	1	468	0.128	0.335	0	1	350	0.026	0.159	0	1
UK firm or industry data	440	0.116	0.320	0	1	468	0.015	0.122	0	1	350	0.143	0.350	0	1
US firm or industry data	440	0.414	0.493	0	1	468	0.327	0.470	0	1	350	0.500	0.501	0	1
Sample related															
Data mid-point after 1980	440	0.589	0.493	0	1	468	0.756	0.430	0	1	350	0.437	0.497	0	1
Firm-level data	440	0.811	0.392	0	1	468	0.887	0.317	0	1	350	0.606	0.489	0	1
R&D-intensive firms/industries	440	0.225	0.418	0	1	468	0.173	0.379	0	1	350	0.037	0.189	0	1
Small firms	440	0.014	0.116	0	1	468	0.017	0.130	0	1	350	0.009	0.092	0	1
Government-funded R&D	440	0.005	0.067	0	1	468	0.006	0.080	0	1	350	0.043	0.203	0	1