The University of Nottingham

research paper series

Globalisation, Productivity and Technology

Research Paper 2011/22

Do Spillovers Matter When Estimating Private Returns to R&D?

by

Markus Eberhardt, Christian Helmers and Hubert Strauss

The Centre acknowledges financial support from The Leverhulme Trust under Programme Grant F/00 114/AM

Leverhulme Centre for Research on Globalisation and Economic Policy

The Authors

Markus Eberhardt is an internal research fellow of GEP, a research associate of the Centre for the Study of African Economies at the University of Oxford, and a Lecturer in the School of Economics, University of Nottingham

Christian Helmers is a research associate of the Centre for the Study of African Economies at the University of Oxford and an Assistant Professor in the Department of Business Administration, Universidad Carlos III, Madrid

Hubert Strauss is a senior economist at the European Investment Bank (EIB) in Luxembourg

Acknowledgements

We are grateful to Stephen Bond, Christine Greenhalgh, Jonathan Haskel, Mark Rogers, Jacques Mairesse, Pierre Mohnen, Ron Smith, Francis Teal, Andrew Toole and Bruno van Pottelsberghe for helpful comments and discussions. The paper was presented at the European Investment Bank, at the University of Oxford, the COINVEST Meeting, the 2010 RES Conference, CISS 2010, the 25th Congress of the EEA, and the Annual Meeting of the German Economic Association 2010 – we thank seminar/session participants for their comments and suggestions. The usual disclaimers apply. Eberhardt gratefully acknowledges financial support from the UK Economic and Social Research Council [grant numbers PTA-031-2004-00345, PTA-026-27-2048] and the George Webb-Medley Endowment Fund, University of Oxford.

Do Spillovers Matter When Estimating Private Returns to R&D?

by

Markus Eberhardt, Christian Helmers and Hubert Strauss

Abstract

A large body of literature estimates private returns to R&D adopting the Griliches knowledge production framework which ignores the potential impact of spillovers on consistent estimation. Using a panel of 12 manufacturing industries across ten OECD economies, we investigate whether ignoring spillovers leads to bias in the estimated private returns to R&D. We compare results from a common factor framework, which accounts for spillovers and other unobserved shocks, to those from a standard Griliches approach. Our findings confirm that conventional estimates conflate own-R&D and spillover effects, implying that spillovers cannot be ignored even when the interest lies exclusively in evaluating private returns to R&D.

JEL classification: D24, L16, C23

Keywords: Productivity, R&D, Spillovers, Common Factor Model

Outline

Non-Technical Summary

Investment in R&D represents one of the few instruments that public policy can affect in the future and due to the particular characteristics of knowledge created through R&D, namely non-excludability and non-exhaustability, private and social returns to R&D generally do not coincide, creating a strong incentive for policy intervention.

There is a vast economic literature attributing an eminent role to R&D in generating productivity gains and long-run growth owing to the generation of spillovers, which account for the difference between social and private returns to R&D. If spillovers are closely linked to R&D, the relevant question is whether the direct effect of R&D on productivity and its direct (i.e., private) returns can be estimated *without also accounting for the spillovers it induces*. The dominating framework in the analysis of private returns to R&D is the Griliches 'knowledge production function', where R&D stock is introduced into a standard Cobb-Douglas production function. While this approach neglects any possibility for spillovers in the empirical framework a separate literature augments the empirical framework with 'spillover variables' to investigate the contribution of spillovers to productivity and thus quantify social returns to R&D investment. These spillover variables are created using a rigid and somewhat ad hoc structure of spillover channels, a practice which reflects the general lack of a clear understanding about the precise channels through which (unobservable) spillovers occur.

This paper asks whether spillovers have to be accounted for within the Griliches knowledge production function framework even when the interest lies exclusively in the estimation of private returns to R&D. If spillovers are unobserved and go unaccounted in the empirical analysis, their presence can lead to correlation between cross-sectional units. Spillovers can, therefore, be regarded as omitted unobserved factors in the error terms. If these unobserved factors are correlated with R&D, the resulting estimates of private returns to R&D are biased and inconsistent. The dedicated knowledge spillover literature implicitly assumes that cross-sectional correlation is *exclusively* generated by R&D spillovers, even though common shocks (e.g. the recent global financial crisis) may equally induce variable correlation across countries. Hence, this approach may fail to produce reliable estimates of private returns in case of empirical misspecification as it may not capture *all* of the cross-sectional dependence in the data. This also implies that a statistically significant spillover variable may not represent genuine knowledge spillovers but rather reflect data dependencies more generally due to a host of other factors common to the countries and industries included in the sample.

In this paper, we adopt a more general 'common factor' framework, which allows us to remain agnostic about the nature and channels of this relationship: our primary interest is in establishing the private returns to R&D investment at the macro-level when accounting for local or global spillovers and common shocks (among other distortions). This means that our results are neither based on ad hoc assumptions about the structure of spillovers nor do we assume that cross-sectional dependence is generated exclusively by knowledge spillovers. To implement our approach empirically, we use an unbalanced panel of ten OECD countries containing data for twelve manufacturing industries covering the period 1980-2005. We compare and contrast the estimates for a Griliches knowledge production function across a number of different specifications with inherently different assumptions about error term independence (lack of R&D and/or other spillover effects) as well as technology homogeneity across countries and/or industries.

Our findings suggest that when spillovers are ignored, private returns to R&D are sizeable; when we account for spillovers of unknown form, which may include other factors than merely R&D spillovers, private returns to R&D are at best modest. In our view, this finding is a strong indication of the presence of spillovers and the indivisibility of R&D from spillovers. Our findings also suggest that commonly employed

R&D spillover variables in the form of some weighted averages of R&D may, on the one hand, fail to adequately capture all of the cross-sectional dependence present in the data and on the other, capture broader cross-sectional data dependencies than solely genuine knowledge spillovers.

"[B]ecause the additive model is not really a very good description of knowledge production, further work on the best way to model the R&D input would be extremely desirable."

Hall, Mairesse, and Mohnen (2009: 33)

1 Introduction

Firms invest in R&D to achieve productivity gains through innovations resulting from their investments.[1] Thus from an aggregate economy perspective, R&D is seen as crucial in achieving productivity growth and has therefore received an enormous amount of attention from policymakers, academics, and the private business sector.[2] As with any type of investment, investment in R&D depends on its expected return — in absolute terms as well as relative to other inputs. In addition, given the particular characteristics of knowledge, non-excludability and non-exhaustability, private and social returns to R&D generally do not coincide. This difference between private and social returns to R&D has motivated a range of policy interventions including direct subsidies and tax credit. From a policy perspective the question of the return to R&D is essential, as R&D spending represents "one of the few variables which public policy can affect in the future" (Griliches, 1979: 115).

Despite the crucial role of investments in R&D, national accounting does not record these in a way that reflects their perceived relevance for productivity growth, although this situation is about to change following an update of the System of National Accounts.[3] But even once R&D is covered in core national accounts, another important issue closely linked to R&D will remain unaccounted for: knowledge spillovers. There is a vast economic literature attributing an eminent role to R&D in generating productivity gains and long-run growth owing to the generation of spillovers (Romer, 1990; Grossman and Helpman, 1991). Notably, spillovers account for the difference between social and private returns to R&D. If spillovers are closely linked to R&D, the relevant question is whether the direct effect of R&D on productivity and its direct (i.e., private) returns can be estimated without also accounting for the spillovers it induces.

Considering the importance of the subject, it is not surprising that there is a substantial number of empirical studies assessing the private and social returns to R&D at the country, regional, industry and firm-level.[4] A closer look at this literature, which is summarized in Table A-I in the Appendix, reveals that the most widely used approach is based on the 'knowledge production function' originally proposed by Griliches (1979). In this approach, R&D stock is added as additional input to a Cobb-Douglas production function. This means that R&D is Hicks-neutral as it shifts the production function without directly affecting returns to the standard inputs, labour and capital. This also implies that R&D enters the production function in an additively separable way, which is a convenient assumption as it allows direct estimation of output elasticities with respect to own-R&D, which are easily converted into returns to R&D.[5] In the Griliches knowledge production function framework, any notion of spillovers is neglected in the empirical specification, a practice maintained in the most recent applications (see for example Doraszelski and Jaumandreu, 2009). In parallel to this approach, there is a large body of research concentrating on the contribution of spillovers to productivity, imposing a rigid structure on the spillover channels in constructing 'spillover variables' based on somewhat *ad hoc* assumptions. This practice reflects the general lack of a clear understanding about the precise channels through which (unobservable) spillovers occur.

This paper asks whether spillovers have to be accounted for within the Griliches knowledge production function framework even when the interest lies exclusively in the estimation of private returns to R&D. If spillovers are unobserved and go unaccounted in the empirical analysis, their presence can lead to correlation between cross-sectional units. Spillovers can, therefore, be regarded as omitted unobserved factors in the error terms. If these unobserved factors are correlated with R&D, the resulting estimates of private returns to R&D are biased and inconsistent.

The dedicated knowledge spillover literature is largely unaware of the econometric importance of accounting for cross-section dependence for consistent estimation and instead concentrates on establishing the impact of 'spillover variables' created in a fashion akin to employing spatial weight matrices.[6] Moreover, this approach implicitly assumes that cross-sectional correlation is exclusively generated by R&D spillovers. Hence, this approach may fail to produce unbiased and consistent estimates of private returns in case of empirical misspecification as it may not capture all of the cross-sectional dependence. This also implies that a statistically significant spillover variable may not represent genuine knowledge spillovers but rather reflect data dependencies more generally due to a host of other factors common to the countries and industries included in the sample.

In this paper, we adopt a more general 'common factor' framework, which allows us to remain agnostic about the nature and channels of this relationship: our primary interest is in establishing the private returns to R&D investment at the macro-level when accounting for any unobserved heterogeneities including local or global spillovers and common shocks. This means that our results are neither based on *ad hoc* assumptions about the structure of spillovers nor do we assume that cross-sectional dependence is generated exclusively by knowledge spillovers. To implement our approach empirically, we use an unbalanced panel of ten OECD countries containing data for twelve manufacturing industries covering the period 1980-2005. We find strong evidence for cross-sectional dependence and the presence of a common factor structure in the data, which we interpret as indicative for the presence of knowledge spillovers *and* additional unobserved cross-sectional dependencies.

We then compare and contrast the estimates for a Griliches knowledge production function across a number of different empirical specifications with inherently different assumptions about error term independence (lack of R&D and/or other spillover effects) as well as technology homogeneity across countries and/or industries. This ensures that our conclusions do not merely reflect specific assumptions imposed on an unknown data generating process.

Our findings suggest that when spillovers in the form of cross-sectional dependence are ignored, private returns to R&D are sizeable; when we account for spillovers of unknown form, which may include other factors than merely R&D spillovers, private returns to R&D are at best modest. In our view, this finding is a strong indication of the presence of spillovers and the indivisibility of R&D from spillovers. If cross-sectional dependence due to knowledge spillovers and/or additional unobserved heterogeneity is present in the data, estimates of the output elasticity with respect to R&D capital confound the direct effect of R&D on output with that of spillovers and a host of other phenomena. Our findings also suggest that commonly employed R&D spillover variables in the form of some weighted averages of R&D may, on the one hand, fail to adequately capture all of the cross-sectional dependence present in the data and on the other, capture broader cross-sectional data dependencies than solely genuine knowledge spillovers.

The remainder of this paper is organized as follows: Section 2 discusses the theory underlying the Griliches knowledge production function at the heart of the literature. Section 3 discusses the theory on knowledge spillovers as well as their empirical measurement. Section 4 introduces the dataset used for our analysis and provides descriptive statistics. Section 5 contains a description of our estimation strategy and Section 6 presents the empirical results. Section 7 concludes.

2 The Knowledge Production Function

The output elasticity with respect to R&D capital, from which the private return to R&D is derived, is commonly estimated adopting a version of the Cobb-Douglas production function framework. Griliches (1979) assumes an augmented production function with value-added Y as a function of standard inputs labour L and tangible capital K as well as 'knowledge capital' R

$$Y = F(L, K, R) \tag{1}$$

With $F(\cdot)$ assumed to be Cobb-Douglas, knowledge capital R is treated as a complement to the standard inputs. According to Griliches, the level of knowledge capital is a function of current and past levels of R&D expenditure

$$R = G[W(B)R\&D] \tag{2}$$

where $W(B)$ is a lag polynomial with B being the lag operator. Equation (2) describes the so-called knowledge production function: the functional relation between knowledge inputs and knowledge output.[7] Griliches then writes (1) as

$$Y = AL^{\alpha}K^{\beta}R^{\gamma}exp^{\lambda t+e} \tag{3}$$

where A is a constant, t is a time index capturing a common linear trend λ and e is a stochastic error term. α, β, γ and λ are parameters to be estimated. Equation (2) can be substituted into Equation (3) to obtain output directly as a function of current and past $R\&D$ expenditure (Hall, 1996). In order to obtain an estimable equation, we take logarithms and use subscripts i and t to denote cross-sectional units and time respectively:

$$y_{it} = \alpha l_{it} + \beta k_{it} + \gamma r_{it} + \lambda_t + \psi_i + e_{it} \tag{4}$$

where lower case letters denote logarithms of the inputs in Equation (3) and λ_t is, more generally than above, a time-specific effect that is (for the sake of exposition) assumed to be common across countries and industries. e_{it} is an error term which contains random shocks to the production and knowledge accumulation processes. Equation (4) contains a measure for R&D capital stock, r_{it}, instead of a lag polynomial of R&D expenditures. We discuss in Appendix B-4 how the R&D capital stock (R) can be constructed from R&D expenditures ($R\&D$). In order to account for cross-section, unit-specific effects that remain constant over time, we also introduce ψ_i. The coefficient γ measures the joint contribution of R&D to productivity and to output prices. γ therefore indicates the elasticity of output with respect to R&D capital, i.e., $\gamma = \frac{\partial Y}{\partial R}\frac{R}{Y}$. Accordingly the gross private rate of return can be obtained as $\rho^G = \gamma\frac{Y}{R}$. Consequently, the net rate of return is $\rho^N = \rho^G - d$ where d is the depreciation rate of R&D capital.

Griliches (1980) noted two important measurement problems with regard to Equation (4): first, conventional measures of capital and labour also contain elements of R&D, which is thus 'double-counted' as R&D workers are included in the total labour force headcount and R&D-related investments in the overall capital stock figure.[8] This was conventionally taken to imply that the coefficient associated with R&D stock is an estimate of the excess gross rate of return to R&D, i.e., the risk premium or supra-normal profit of R&D investment over other investment. Second, since R&D is treated as an 'intermediate expense' in the calculation of value-added, measured value-added is too small by that amount.

Schankerman (1981) discusses the distorting impact of these mismeasurements in both a growth accounting and regression framework. Within the confines of the latter he notes that the failure to recognise the 'double-counting' of R&D inputs and the 'expensing' of R&D can be framed as an omitted variable problem. He goes on to show that the omission of the share of R&D workers in total labour and of R&D-related investments in total investment leads to a downward bias on the R&D stock coefficient, which cannot be interpreted as "an excess return in any simple sense" (Schankerman, 1981:456). The 'expensing bias' resulting from the failure to account for R&D intensity may be either positive or negative, such that the sign of the combined bias is *a priori* ambiguous. Some of the existing empirical evidence in cross-section data suggests an overall downward bias in the coefficient of the R&D stock (Schankerman, 1981; Hall and Mairesse, 1995) although the significance of this bias in panel datasets accounting for fixed effects is subject to some debate (Cuneo and Mairesse, 1984; Hall and Mairesse, 1995; Guellec and van Pottelsberghe, 2004). Our strategy to deal with these econometric difficulties will be twofold: firstly, we show that the unobserved common factor model adopted in our empirics and detailed in Section 3.2.2 is theoretically appropriate to tackle the excess returns and expensing biases. Secondly, we follow the suggestion by Schankerman (1981) and investigate the significance of these biases in our data using both adjusted input values to account for 'double-counting' and augmented empirical equations to account for 'expensing' of R&D, with results discussed briefly in Section 6 and presented in more detail in a Technical Appendix.

The overall validity of the Griliches knowledge production function approach rests on the assumption of perfectly competitive factor markets, full capacity utilization, as well as the absence of spillover effects — the latter is econometrically represented by the cross-sectional independence of error terms e_{it} in Equation (4). While implied by our notation in the empirical setup described above, there is no obvious reason to require the input coefficients of the knowledge production function to be the same across countries or industries ($\alpha_i = \alpha$, $\beta_i = \beta$, $\gamma_i = \gamma$).[9] We investigate these issues in greater detail in the following.

3 Knowledge Spillovers and other Cross-Section Dependencies

In this section we introduce a second empirical literature that extends the Griliches knowledge production framework to measure productivity gains that arise from R&D spillovers. We discuss the main assumptions routinely made in this literature, prime amongst which is the specification of a known, additively separable, functional form which allows the estimation of separate coefficients associated with own-R&D and R&D spillovers respectively. The approach rests on the assumption that any cross-sectional dependence present in the data reflects R&D spillovers and that these are accurately captured by the coefficient associated with the spillover variable. In order to provide an answer to our research question — "Do Spillovers Matter When Estimating Private Returns to R&D?" — that is not dependent on such *ad hoc* assumptions, we then introduce a more flexible encompassing empirical framework.

3.1 Knowledge Spillovers

Arrow (1962) pointed out that knowledge is distinct from the traditional production factors labour and physical capital. The distinguishing features are (i) non-excludability, and (ii) non-rivalry of knowledge. These features lead to the fact that "we do not deal with one closed industry, but with a whole array of firms and industries which borrow different amounts of knowledge from different sources according to their economic and technological distance from them" (Griliches, 1979:103). Hence, knowledge spills over to other actors which do not pay the full cost of accessing and using the knowledge. The process of unintentional knowledge transmission from one actor to another is commonly referred to as 'knowledge spillovers'.[10] This implies that the return on investment in knowledge is partly private and partly public (Keller, 2004).

3.2 Spillovers in the Knowledge Production Function

3.2.1 Standard Approaches

Given the fundamentally unobservable nature of knowledge spillovers, directly quantifying their magnitude is a difficult task. Within the production function framework, the most common approach in the literature proceeds in two steps — we assume $i = 1, \dots, N$ industries within a single country for simplicity of exposition. First, TFP is estimated or computed from value-added and standard factor inputs labour and physical capital; in a second step the resulting TFP estimates are regressed on an industry's own R&D and some measure of knowledge spillovers:

$$\text{TFP}_{it} = g\left(R_{it}, \sum_{k}^{N} \omega_k R_{kt}\right) \tag{5}$$

where R_{it} denotes the R&D stock of industry i and the second term in parentheses captures spillovers received from all other industries ($i \neq k$), with ω_k some explicit weights structuring the relative 'importance' of industries. This setup allows for a differential impact of other industries' R&D stocks on industry i's productivity but comes at the cost of a rigid structure in the specification of ω_k, usually based on somewhat *ad hoc* assumptions. Examples of the imposed structure for spillovers include Input/Output tables (Goto and Suzuki, 1989; Keller, 2002a), import weights (Coe and Helpman, 1995; Keller, 1998), inward/outward FDI or shares of foreign affiliates' sales in domestic sales of an industry (van Pottelsberghe and Lichtenberg, 2001; Baldwin et al., 2005), geographic distance (Keller, 2002b), distance to technology frontier as measured by TFP differences (Griffith et al., 2004; Cameron et al., 2005; Acemoglu et al., 2006), and measures of technological proximity (Conley and Ligon, 2002; Guellec and van Pottelsberghe, 2004).[11]

Equation (5) can be estimated as

$$\text{tfp}_{it} = \psi_i + \gamma r_{it} + \chi \sum_{k=1}^{N} \omega_k r_{kt} + \varepsilon_{it} \tag{6}$$

where lower case letters denote logarithms and ε_{it} is a stochastic shock. Equation (6) is commonly augmented with time dummies to purge additional correlation across industries, arising from common shocks (recessions, policy changes) which affect all industries in the same way. If the sample contains industry-level data from several countries, the specification usually also includes country fixed effects to capture country-specific effects.

The underlying assumptions made in this setup are worth emphasizing: Equation (6) assumes that spillovers affect TFP linearly as captured by the corresponding parameter χ. The spillover effect is additively separable from the own-R&D effect γ. More importantly, the model suggests that industry TFP levels are correlated exclusively because of R&D spillovers and that the spillover measure captures the nature of these spillovers appropriately, that is, conditional on $\sum_{k=1}^{N} \omega_k r_{kt}$, the residuals ε_{it} are cross-sectionally independent. Furthermore, with special reference to the analysis of industry- or country-level data with a substantial time-horizon, it is also assumed that the empirical specification captures the long-run equilibrium relationship and is not distorted by dynamic misspecification or neglect of salient time-series properties of the data. Econometrically, these assumptions translate into well-behaved, serially uncorrelated, stationary and cross-sectionally independent regression residuals $\hat{\varepsilon}_{it}$.

In order to avoid empirical restrictions based on *ad hoc* assumptions about the nature of spillover channels as well as all of the other concerns raised above, we suggest an empirical strategy which (i) can capture knowledge spillovers of unknown form together with any other unobserved heterogeneities that may cause cross-sectional correlation; (ii) allows for heterogeneous production technology across industries; and (iii) is concerned with the appropriate treatment of dynamics and time-series properties more generally.

3.2.2 Unobserved Common Factor Framework

The common factor approach assumes that the error term as well as the covariates in the empirical model contain a finite number of unobserved common processes ('factors'), whose impact may differ across industries or countries. Recent work in this area has emphasised the distinction between 'strong' factors representing global shocks such as the recent global financial crisis, and 'weak' factors such as spillovers between a limited group of industries or countries (Holly et al., 2010; Chudik et al., 2011). This setup has particular appeal for the present analysis of returns to own-R&D in a set of interconnected OECD countries that are subject to common shocks which, however, may impact individual economies differentially, and where R&D may spill over from one industry or economy to another following a complex, unknown, and non-symmetric structure.

We can illustrate the model setup in a simplified version of Equation (4) with a single input x_{it} and (for generality) heterogeneous technology parameter $\beta_i = \beta + \varpi_i$ where $\varpi_i \sim iid(0, \sigma^2_{\varpi})$

$$y_{it} = \beta_i x_{it} + u_{it} \tag{7}$$

Cross-sectional dependence arises from the multi-factor error structure and the assumed driving force of the input

$$u_{it} = \varphi_i f_t + \psi_i + \varepsilon_{it} \tag{8a}$$

$$x_{it} = \varrho_i f_t + \pi_i g_t + \phi_i + e_{it} \tag{8b}$$

where e_{it} and ε_{it} are stochastic shocks. The setup assumes that latent processes drive both productivity as well as the inputs, albeit not necessarily with the same strength ('factor loadings' φ_i and ϱ_i differ from each other). The fact that the regressor as well as the error term share a common factor f_t implies that if the factor loadings φ_i and ϱ_i are on average nonzero, estimating (7) without accounting for f_t produces biased and inconsistent estimates of $E[\beta_i] = \beta$, as can be shown by simple substitution

$$y_{it} = \underbrace{\left(\beta_i + \varphi_i \varrho_i^{-1}\right)}_{\zeta_i} x_{it} + \underbrace{\psi_i - \varphi_i \varrho_i^{-1} \phi_i}_{\eta_i} + \underbrace{\varepsilon_{it} - \varphi_i \varrho_i^{-1} \pi_i g_t - \varphi_i \varrho_i^{-1} e_{it}}_{\varsigma_{it}} \tag{9}$$

$$= \zeta_i x_{it} + \eta_i + \varsigma_{it}$$

This idea extends to multiple factors and the multivariate context, such as the Griliches knowledge production function where the main focus is on the coefficient of own-R&D: if the unobservable f_t is merely a 'weak' factor (representing local spillovers between a small number of industries) then the estimate of the β coefficient may not be seriously biased; however, if we have multiple factors of the 'weak' and 'strong' type, the β coefficient is not identified.[12]

As suggested above, the common factor framework can also account for the omitted variable bias arising from double-counting and expensing of R&D (Schankerman, 1981). If observed labour, capital stock, and value-added are 'mismeasured' by the share of R&D workers in total labour (s_{it}), the share of R&D capital in capital (δ_{it}), and the measured R&D intensity (θ_{it}) respectively, then the true relationship can be represented in a variant of Equation (4) (adapted from Equation (10) in Schankerman, 1981) as

$$y_{it} = \alpha(l_{it} - s_{it}) + \beta(k_{it} - \delta_{it}) + \gamma r_{it} - \theta_{it} + \lambda_t + \psi_i + e_{it} \tag{10}$$

$$= \alpha l_{it} + \beta k_{it} + \gamma r_{it} + \left[\lambda_t + \psi_i - \alpha s_{it} - \beta \delta_{it} - \theta_{it}\right] + e_{it} \tag{11}$$

where λ_t and ψ_i are time- and country-industry specific effects. Provided the omitted shares (s, δ) and R&D intensity (θ) each display some commonalities across a subset of country-industries, e.g. increase over time in all R&D-intensive industries or increase within all industries of one country, the omitted variables in brackets can be represented by a combination of unobserved common factors (here for simplicity: h_t, i_t and j_t) with heterogeneous factor loadings (and a set of intercept terms)

$$y_{it} = \alpha l_{it} + \beta k_{it} + \gamma r_{it} + \left[\pi_{1i} h_t + \pi_{2i} i_t + \pi_{3i} j_t + \pi_{4i}\right] + e_{it} \tag{12}$$

Since these common factors are correlated with the R&D stock (Schankerman, 1981:456), failure to account for their presence leads to the identification problem highlighted above. The omitted variable problem described as the source of the R&D 'double-counting' and 'expensing' bias can thus be accommodated econometrically in our encompassing empirical framework. We will nevertheless also estimate a version of Equation (10) in which we use observed s_{it}, δ_{it}, and θ_{it} to account for both 'double-counting bias' and 'expensing bias' (see Section 6.1).

4 Data

The dataset comprises information on up to twelve manufacturing industries (SIC 15-37 excluding SIC 23)[13] in ten countries (Denmark, Finland, Germany, Italy, Japan, Netherlands, Portugal, Sweden, United Kingdom, and the US) over a time period of up to 26 years from 1980 to 2005, yielding a total of 2,637 observations — see Tables 1 and 2 for details.[14] All of the results presented assume the country-industry as the unit of analysis (panel group member i), of which we have $N = 119$, yielding an average $T = 22.2$ time-series observations per industrial sector. The data are taken from a number sources including the EU KLEMS dataset for the production data, the OECD for R&D expenditure and Eurostat and the OECD for GDP deflators.

All monetary variables in our dataset are expressed in million Euros and deflated to 1995 price levels using either country- or industry-level deflators. We use double-deflated value-added, total number of hours worked by persons engaged and total tangible assets by book value as our measures of output, labour and capital stock respectively. R&D stock is taken from KLEMS and extended to 2004 and 2005 using OECD data. In addition we construct the R&D capital stock series for Portugal, following the method adopted by KLEMS. We provide more details on data construction and assumptions made in Appendix B.

Table 3 contains descriptive statistics for the data sample used in our regression analysis. In Figure 1 we provide box plots for value-added, physical capital stock and R&D capital stock for the year 2005 (all deflated by million working hours), sorted by median value. As can be seen in the cross-country analysis of the left column, Japan is near the top for all three measures, whereas Portugal maintains the bottom position. The latter country aside, the distribution of value-added and physical capital stock per hour worked is relatively similar across these economies and has a narrow interquartile range, whereas the R&D capital stock per hour worked varies much more substantially. For the cross-industry analysis in the right column we can note that the Chemicals industry (SIC 24) tops all three graphs while Textiles (SIC 17-19) and Other Manufactures (SIC 36/37) can be found toward the bottom. Cross-industry variation is much stronger than cross-country variation and features more outliers, particularly for the R&D stock variable.

As a means of pre-estimation analysis of the data, we investigate the time-series and cross-section properties of all variables using panel unit root tests of the first (Maddala and Wu, 1999) and second generation (Pesaran, 2007), average cross-section correlation coefficients as well as a formal test for cross-section dependence by Pesaran (2004). Detailed results are presented in a Technical Appendix. We also employ these tests in our residual diagnostics for each of the empirical models presented below. The panel unit root tests suggest that all variables are integrated of order one. The analysis of cross-section correlation indicates substantial dependence for the variables in levels as well as first differences.[15]

5 Estimation Strategy

By the nature of our research question, the empirical implementation will be carried out using different estimators, each of which will impose different assumptions about the underlying data generating process, which can in part be tested using a range of diagnostic tests applied to the residuals.[16] This ensures that our empirical findings do not simply mirror specific assumptions imposed by different empirical specifications and estimators. We employ the following general regression equation and use the scheme in Table 4 to structure the different approaches into a common framework.

$$\begin{aligned} y_{it} &= \alpha_i l_{it} + \beta_i k_{it} + \gamma_i r_{it} + \lambda_{it} + \psi_i + e_{it} \\ & e_{it} = \rho_i e_{i,t-1} + u_{it} \end{aligned} \tag{13}$$

where l, k and r are labour, capital stock and R&D stock (in logarithms).

A first distinction is to be made between common and heterogeneous parameter models: the former, 'pooled' estimators, assume common technology parameters on factor inputs across all countries and industries ($\alpha_i = \alpha$, $\beta_i = \beta$, $\gamma_i = \gamma \, \forall \, i$), while the latter relax this assumption to a varying degree.[17] Typical pooled estimators include the least squares estimator augmented with year dummies (POLS) or the Two-way Fixed Effects estimator (2FE) which contains country-industry as well as time fixed effects. 'Mean Group' type estimators allow for technology heterogeneity by running country-industry specific regressions and then averaging the coefficients across the panel. Results for individual country-industry pairs are unreliable (unless T is large) and are often difficult to interpret, whereas panel averages establish a reliable mean estimate (Boyd and Smith, 2002). In Table 4 the distinction between common and heterogeneous technology parameters is between the upper and lower panels.

A second distinction is made between static and dynamic models, which is implemented for the common and heterogeneous technology models, respectively. Investigating long-run equilibrium relations in a static model without any lagged variables may oversimplify the dynamic adjustment of the system and may mistake short-run deviations for long-run effects. A first attempt at dealing with this is to specify a simple Autoregressive Distributed Lag model (ARDL), which can be derived from Equation (13) for $\rho_i \neq 0$

$$\begin{aligned} y_{it} &= \rho_i y_{i,t-1} + \alpha_i l_{it} - \rho_i \alpha_i l_{i,t-1} + \beta_i k_{it} - \rho_i \beta_i k_{it} + \gamma_i r_{it} - \rho_i \gamma_i r_{it} \\ &\quad + (\lambda_{it} - \rho_i \lambda_{i,t-1}) + (1 - \rho_i)\psi_i + u_{it} \end{aligned} \tag{14}$$

Equation (14) is commonly estimated in an unrestricted version without the non-linear ('common factor') restrictions implied. Based on empirical testing, the long-run relationship in the data can then be evaluated either with or without restrictions. Apart from standard pooled estimators (POLS, 2FE) we also employ the dynamic micro-panel estimator by Blundell and Bond (1998, BB).[18] The latter deals with the problem of 'Nickell bias' (Nickell, 1981) in a dynamic panel data model with fixed effects, which yields inconsistent estimates in samples with limited T. The unique instrumentation employed, using transformed equations and lagged values of endogenous variables, has the additional attraction that it can provide 'internal' instruments for any endogenous variable in the model. Despite a number of problems (Bowsher, 2002; Roodman, 2009) this type of micro panel estimator has become very popular for use in macro panel data. The BB estimators solves the identification problem we discussed in Section 3.2.2 (correlation between observed inputs and unobservables/TFP) but relies on the crucial assumption that technology parameters (α, β, γ) do not differ across country-industries.[19] The distinction between static and dynamic models is highlighted in Table 4.

A third distinction relates to the concerns over cross-section dependence, including both knowledge spillovers as well as any other type of spillovers and/or common shocks. As we developed above, the various types of cross-section correlation are modelled in our empirical strategy using unobserved common factors.[20] The distinction between the left and right column in Table 4 represents different assumptions about the impact of these unobservables.

All the pooled estimators in the left column of Table 4 are augmented with year dummies (in the 2FE case implicitly) which can account for the presence of unobserved common factors provided their impact does not differ across country-industries. For the empirical models in Equations (13) and (14) this would imply $\lambda_{it} = \lambda_t$. The evolution of the unobservables over time is not constrained in any way, thus could be linear or nonlinear, stationary or nonstationary. In the lower panel of the Table, the Mean Group estimator with variables in deviation from the cross-section mean (CDMG) maintains the same assumption about a common impact of unobservables across country-industries but allows for differential technology parameters.

The right column of Table 4 contains estimators which allow for the impact of unobserved common factors to differ across countries and industries. Among the Mean Group type estimators in the lower panel of the Table, the Pesaran and Smith (1995) Mean Group (MG) estimator can be augmented with country-industry specific linear trends which allow for a differential impact of unobservables across country-industries whilst imposing linearity on their evolution. The Pesaran (2006) Common Correlated Effects (Pooled or Mean Group) estimators account for unobserved common factors with heterogeneous factor loadings by using cross-section averages of the dependent and independent variables as additional regressors. This allows for more flexibility as the impact of the unobserved common factors can differ across country-industries while the evolution of these factors may be nonlinear or even nonstationary (Kapetanios et al., 2011).[21] To see the intuition behind this approach, consider the cross-section average of our pet model from Equations (7), (8a) and (8b) above, replicated here for convenience[22]

$$\begin{aligned} y_{it} &= \beta_i x_{it} + \varphi_i f_t + \psi_i + \varepsilon_{it} && (15)\\ \bar{y}_t &= \bar{\beta}\bar{x}_t + \bar{\varphi} f_t + \bar{\psi} \qquad \text{given } \bar{\varepsilon}_t \to 0 \text{ as } N \to \infty \\ \Leftrightarrow \quad f_t &= \bar{\varphi}^{-1}(\bar{y}_t - \bar{\psi} - \bar{\beta}\bar{x}_t) && (16) \end{aligned}$$

where cross-section averages at time t are defined as $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$ and $\bar{x}_t = N^{-1}\sum_{i=1}^{N} x_{it}$.[23] In words, as the cross-section dimension becomes large the unobserved common factor f_t can be captured by a combination of cross-sectional averages of y and x. Substitution for f_t in Equation (15) yields

$$\begin{aligned} y_{it} &= \beta_i x_{it} + \varphi_i \bar{\varphi}^{-1}\left(\bar{y}_t - \bar{\psi} - \bar{\beta}\bar{x}_t\right) + \psi_i + \varepsilon_{it} && (17)\\ \Leftrightarrow \quad y_{it} &= \beta_i x_{it} + \pi_{1i}\bar{y}_t + \pi_{2i}\bar{x}_t + \pi_{3i} + \varepsilon_{it} && (18) \end{aligned}$$

As can be seen the parameters of $\bar{y}_t$ and $\bar{x}_t$ as well as the intercept π_{3i} must be country-industry specific to capture the heterogeneity in the factor loadings φ_i. In the heterogeneous technology version of the estimator (CMG), where we allow for $\beta_i \neq \beta$, this is achieved by construction since each country-industry is estimated separately. In the pooled version (CCEP) the cross-section averages need to be interacted with country-industry dummies, so that each country-industry can have a different parameter on the cross-section averages. Both estimators can accommodate a fixed number of 'strong' common factors and an infinite number of 'weak' common factors (Chudik et al., 2011), where the former can be thought of as common global shocks and the latter as local/regional spillover effects. The focus of this estimation approach is to obtain unbiased estimates for β or the mean of the heterogeneous β_i; since various averages of the unknown parameters are contained in π_{1i}, π_{2i} and π_{3i} these cannot be interpreted and should be seen as merely accounting for the cross-section dependence in the data.

6 Results

In the following we discuss the empirical results from our study of ten OECD economies with up to twelve manufacturing sectors each. We follow the scheme in Table 4, beginning with common technology models (static, dynamic), then moving on to heterogeneous technology models (static, dynamic). Within each of these four groups, estimators differ in their assumptions about cross-section dependence/common factors. In order to evaluate rival empirical models we use a number of diagnostic tests including a Wald test of constant returns to scale ($\alpha + \beta + \gamma = 1$), serial correlation tests (in the static models only), common factor restriction tests (in the dynamic models only), residual cross-section correlation tests (Pesaran, 2004) and residual stationarity tests (Pesaran, 2007). In addition we provide the root mean squared error (RMSE) statistic for each regression model to indicate a measure for goodness of fit.

6.1 Common Parameter Models

Table 5 contains the results for standard pooled panel estimators in their *static* specification (POLS, 2FE, FD) as well as for the CCEP estimator in its standard version and augmented with common year dummies. All five models yield statistically significant and sensible parameter estimates for capital and labour inputs, ranging from .2 to .5 and .45 to .65 respectively. The coefficient of the R&D stock is large and highly significant in the POLS case and to a lesser extent in the 2FE and CCEP models. Although of relatively similar magnitude, the R&D coefficient in the FD model is not significant at the 5% level. All parameter estimates are economically plausible.

Turning to the diagnostics, it is suggested that POLS and 2FE yield nonstationary residuals and we therefore cannot rule out spurious results, even in a panel regression (Kao, 1999). Serial correlation is present in all five models (AR(1) is to be expected in the FD case) and curiously the residual CD tests for cross-section independence seem to reject in case of CCEP estimators. The measure of fit indicates that the FD and CCEP models have similar residual standard deviations, which are much smaller than those for the POLS and 2FE models.

Our interpretation of these results is that the standard pooled models in levels (POLS, 2FE) are seriously misspecified, given their serially correlated and nonstationary residuals. Since these models do not seem to suffer from cross-sectionally correlated residuals and the FD yields more favourable diagnostics we suggest that the source of the misspecification derives either from the (lack of) dynamics or the erroneous pooling of all country-industries (common technology). The CCEP models fail to address the concerns for which they were developed, namely to account for all cross-section dependencies; again, possible causes include the two misspecifications suggested. Our preferred pooled model in the static specification is thus the FD, which yields an R&D coefficient roughly one half in magnitude of the standard OLS estimator, albeit statistically insignificant.

Table 6 turns to the results for the *dynamic* specifications. In order to ease comparison with the static results we only report the long-run coefficients implied by the common factor restrictions (ARDL model estimates based on Equation (14) are available on request). Implied long-run coefficients for capital and labour vary substantially across the five models presented, from .1 to .9 and -.5 to .7, respectively. All but the POLS model in [1] result in very low and/or statistically insignificant R&D capital. For the POLS estimator it seems that identification of capital stock in the presence of R&D stock is challenging and although the diagnostic tests indicate some favourable residual diagnostics these results are still somewhat questionable — the identification problem highlighted in Equation (9) above is most likely the culprit for this outcome. The poor performance of the BB estimator (negative albeit insignificant labour coefficient), relying on lagged levels variables as instruments for contemporaneous first differences and on lagged differences for levels, highlights the persistence and likely nonstationarity of the data. The two CCEP estimators yield similar results, with R&D capital insignificant and around .03.

Diagnostics for these models seem to suggest that only the 2FE and CCEP models yield stationary residuals, while the popular micro-panel estimator (BB) further fails the instrument validity (Sargan) test.[24] Once we take the possibility of cross-section dependence (spillovers, common shocks) explicitly into account in models [4] and [5] we see a substantial reduction in the coefficient of R&D capital, and we can no longer detect a statistically significant impact. Given their favourable diagnostics, our preferred dynamic pooled models are the standard and augmented CCEP in columns [4] and [5].

We have argued in Section 3.2.2 above that the concerns over double-counting and expensing of R&D should be alleviated in a panel model accounting for unobserved common factors. We nevertheless also offer results that are obtained from explicitly correcting the input variables and value-added for mis-measurement following Schankerman (1981). However, the data required to correct for double-counting and expensing are only available for a subset of countries, industries and time periods. Hence, the sample used to explore the effect of explicitly correcting the data is less than 30% of the size of the original sample. This lack of data allows us only to implement the static specification of our pooled model for which we estimate two specifications: i) directly correcting the input variables and ii) augmenting the specification with the omitted variables. Furthermore, the CCEP estimators were dropped since their use would have lead to a further halving of the sample while it is also unlikely that these estimators would perform as expected in the resulting short-T panel (average $T = 7.5$). To briefly summarize, we find that the results obtained from the corrected data suggest some downward bias in the R&D coefficient, mostly due to double-counting, but produce statistically insignificant R&D coefficients (except for POLS). The models which add s, δ and θ to the regression show very little impact on the R&D capital coefficient throughout. A more detailed discussion of the approach and the corresponding results is relegated to a Technical Appendix.

6.2 Heterogeneous Parameter Models

In our results for the static and dynamic models in Tables 7 and 8 we focus on the most flexible specification where each country-industry is allowed to follow a different production function — we also investigated intermediate models using country- or industry-level regressions which yielded qualitatively similar results regarding R&D capital stock (available on request).

The average labour coefficients in our *static* results in Table 7 are again quite similar across all models, between .56 and .70 and thus close to the macroeconomic data on factor income-share in developed economies (e.g. Gomme and Rupert, 2004). Capital coefficients are however notoriously difficult to estimate precisely, so it is not surprising that only in the CDMG model in column [2] we obtain statistically significant results. Results for the capital coefficient in the two CMG specifications in columns [3] and [4] are plausible (given the imprecision) if somewhat on the low side. Only the CDMG model yields a statistically significant R&D stock coefficient and it bears noting that overall the CDMG results are very similar to those of the pooled OLS model in Table 5.

Once we take the diagnostic tests into account, we can see that MG and CDMG suffer from cross-sectionally dependent, serially dependent and possibly nonstationary residuals — provided we want to distinguish between empirical models using these testing procedures the conclusion must be that these models are seriously misspecified. The two CMG models obtain much more favourable diagnostic results, no longer rejecting cross-section independence, with the model without country trends in column [3] being preferable due to the more convincing evidence for residual stationarity.

We can conclude from this analysis that the imposition of a rigid structure on the nature of spillovers and common shocks — as is the case in the CDMG model where shocks are assumed to impact all country-industries in an identical way — produces a spuriously high coefficient of the R&D capital stock, which is substantially reduced once we allow for a more flexible structure in the CMG models.[25]

The *dynamic* models for which we present results in Table 8 (based on empirical testing we impose common factor restrictions; full results are available on request) represent a considerable challenge for our data given the moderate time-series dimension available: these models are estimated with between 8 and 17 covariates in the CDMG and trend-augmented CMG models respectively. Due to this dimensionality problem we are forced to drop a number of countries (GER, PRT, SWE) from the analysis in the CMG models — results for the MG and CDMG in this reduced sample were qualitatively very similar to those presented so we report results for the larger sample for these two models. Given these data problems we only view these results as tentative evidence and merely highlight the similar patterns to the static heterogeneous models discussed above: CDMG yields a spuriously high R&D coefficient due to the imposition of common impact of unobservables across country-industries; once this assumption is relaxed in the CMG models in columns [3] and [4] the coefficient drops substantially in magnitude and is no longer statistically significant.[26] Diagnostic tests again suggest that MG and CDMG yield possibly nonstationary residuals and all models raise some concerns over residual cross-section dependence.

In summary, our empirics have paid particular attention to residual cross-section dependence, which in economic terms can be interpreted as knowledge spillovers and/or other unobserved shocks but econometrically raises serious concerns regarding consistency of the regression estimates. We offer a number of alternative specifications for the empirical model, allowing for dynamics as well as technology heterogeneity across countries. We find across these alternatives that models which yield a large and statistically significant coefficient of own-R&D are seriously misspecified (nonstationary, serially correlated and/or cross-sectionally dependent residuals). In contrast, once our diagnostic tests are more favourable the coefficient of own-R&D always drops considerably and becomes statistically insignificant. We take this as a clear indication that spillovers, be they true knowledge spillovers or other common shocks, matter and cannot be ignored even when the interest lies exclusively in estimating private returns to R&D.

7 Concluding Remarks

In this study we asked whether returns to R&D can be estimated in a standard Griliches-type production function framework ignoring the potential presence of knowledge spillovers between cross-sectional units as well as other cross-section dependencies. Finding an answer to this question is relevant considering the vast amount of empirical work either implementing a Griliches-type production function under the assumption of cross-section independence, or investigating knowledge spillovers, assuming a known, additively separable functional form for R&D and spillovers and positing that no other cross-section dependencies are captured by the 'R&D spillover' variable. The main claim of this paper is that the Griliches framework is inadequate even when the analysis focuses exclusively on private returns to R&D.

Using data for 12 industrial sectors in 10 OECD countries our results suggest the conventional Griliches-type knowledge production function model is indeed seriously misspecified, with diagnostic tests pointing at nonstationary and serially correlated residuals. Across static and dynamic as well as pooled and heterogeneous parameter models we can trace a pattern whereby estimators which explicitly account for cross-section dependence and are robust to variable nonstationarity yield substantially lower coefficients for the R&D capital stock which are statistically insignificant in most cases. These findings suggest that conventional approaches imply large and significant private returns to R&D, while specifications accounting for cross-section dependence imply relatively limited private returns to R&D.

These results may be explained by at least two types of arguments and most likely by a combination of the two. First, R&D is a worthwhile undertaking. Yet, its value stems from a complex mix of own R&D successes and spillovers received rather than from a clearly identifiable stream of returns to an industry's own R&D investment: once we account for spillovers, private returns to R&D are modest. Second, the empirical approach taken here does not just account for knowledge spillovers but for any other cross-section dependencies, including other types of productivity spillovers unrelated to R&D as well as the impact of common shocks. The true, social return to R&D investment is likely to be substantially higher. It is partly the result of interactions between factor inputs as well as between countries and industries. It can therefore not be extracted in a *ceteris paribus* fashion as is common in a knowledge production function building on additive separability and focusing on private returns.

Our analysis, therefore, offers two conclusions: first, even when the objective is to identify only private returns to R&D, spillovers cannot be ignored. Second, only including measures capturing R&D spillovers in the empirical equation is unlikely to account appropriately for cross-sectional dependence that is commonly generated by the complex interplay of a range of unobserved processes. Instead, the coefficient associated with the 'R&D spillover' variable is likely to at least in part capture common shocks and cross-sectional dependence that arises for reasons other than genuine knowledge spillovers. The common factor approach adopted in our analysis offers a way of recovering private returns by stripping the estimates from any other confounding factors.

While our analysis sheds some light on the importance of spillovers and other causes of cross-section correlation in the estimation of private returns to R&D, we do not recover a parameter associated with spillovers and therefore cannot make any statements regarding the social returns to R&D. If social returns are the object of interest, more structure needs to be imposed on the nature of spillovers to be able to recover the corresponding parameter within a spatial econometric framework. Any such analysis thus necessarily involves the question of how to measure spillovers.[27] We deliberately avoided addressing this question by adopting an agnostic common factor approach in order to escape the need to make *ad hoc* assumptions about the unobserved structure of spillovers. In our mind, the search for a more appropriate specification of the knowledge production function that accounts for the true nature of cross-sectional interdependencies and allows identification of private and social returns to R&D should be regarded as the main challenge for the investigation of returns to R&D in years to come.

References

[1] Acemoglu, Daron, Philippe Aghion and Fabrizio Zilibotti (2003): 'Distance to Frontier, Selection, and Economic Growth', *Journal of the European Economic Association* Vol. 4, No. 1, pp. 37-74.

[2] Arellano, Manuel and Stephen R. Bond (1991): 'Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations', *Review of Economic Studies* Vol. 58, No. 2, pp. 277-297.

[3] Arrow, Kenneth (1962): 'Economic Welfare and the Allocation of Resources for Inventions', in: R. Nelson (ed.) *The Rate and Direction of Inventive Activity*, Princeton University Press.

[4] Azariadis, Costas and Allan Drazen (1990): 'Threshold Externalities in Economic Development', *Quarterly Journal of Economics* Vol. 105, No. 2, pp. 501-26.

[5] Baldwin, Richard, Henrik Braconier and Rikard Forslid (2005): 'Multinationals, Endogenous Growth, and Technological Spillovers: Theory and Evidence', *Review of International Economics* Vol. 13, No. 5, pp. 945-963.

[6] Banerjee, Abhijit V. and Andrew F. Newman (1993): 'Occupational Choice and the Process of Development', *Journal of Political Economy* Vol. 101, No. 2, pp. 274-98.

[7] Blundell, Richard and Stephen R. Bond (1998): 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics* Vol. 87, No. 1, pp. 115-143.

[8] Bowsher, Clive (2002): 'On testing overidentifying restrictions in dynamic panel data models', *Economics Letters* Vol. 77, pp. 211-220.

[9] Boyd, Derick and Ron P. Smith (2002): 'Some econometric issues in measuring the monetary transmission mechanism, with an application to developing countries', in: Mahadeva, Lavan and Peter Sinclair (eds) *Monetary Transmission in Diverse Economies*, Cambridge University Press, pp. 68-99.

[10] Braconier, Henrik and Fredrik Sjöholm (1998): 'National and international spillovers from R&D: Comparing a neoclassical and an endogenous growth approach', *Review of World Economics* Vol. 134, No. 4, pp. 638-663.

[11] Cameron, Gavin (1996): 'Innovation and Economic Growth', Centre for Economic Performance Discussion Paper No. 277, London School of Economics, February 1996.

[12] Cameron, Gavin, James Proudman, and Stephen Redding (2005): 'Technological convergence, R&D, trade and productivity growth', *European Economic Review* Vol. 49, No. 3, pp. 775-807.

[13] Chudik, Alexander, M. Hashem Pesaran and Elisa Tosetti (2011): 'Weak and Strong Cross Section Dependence and Estimation of Large Panels', *Econometrics Journal* Vol. 14, No. 1, pp. C45-C90.

[14] Coe, David and Helpman Elhanan (1995): 'International R&D Spillovers', *European Economic Review* Vol. 39, No. 5, pp. 859-887.

[15] Coe, David, Elhanan Helpman and Alexander Hoffmaister (2009): 'International R&D Spillovers and Institutions', *European Economic Review* Vol. 53, No. 7, pp. 723-741.

[16] Conley, Timothy and Ethan Ligon (2002): 'Economic Distance and Cross-Country Spillovers', *Journal of Economic Growth* Vol. 7, No. 2, pp. 157-187.

[17] Crepon, Bruno, Emmanuel Duguet and Jacques Mairesse (1998): 'Research, Innovation and Productivity: An Econometric Analysis at the Firm Level', *Economics of Innovation and New Technology* Vol. 7, No. 2, pp. 115-158.

[18] Cuneo, Philippe and Jacques 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, pp. 376-392.

[19] Doraszelski, Ulrich and Jordi Jaumandreu (2008): 'R&D and Productivity: Estimating Production Functions when Productivity is Endogenous', CEPR Discussion Papers No. 6636, January 2008.

[20] Eberhardt, Markus and Christian Helmers (2010): 'Untested Assumptions and Data Slicing: A Critical Review of Firm-Level Production Function Estimators', Oxford University, Department of Economics Discussion Paper Series No. 513, November 2010.

[21] Eberhardt, Markus and Francis Teal (2011): 'Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics', *Journal of Economic Surveys* Vol. 25, No. 1, pp. 109-155.

[22] Edworthy, Emma and Gavin Wallis (2007): 'Treating Research and Development as a Capital Asset', *Economic & Labour Market Review* Vol. 1, No. 2, pp. 16-25.

[23] Gomme, Paul and Peter Rupert (2004): 'Measuring Labor's Share of Income', Policy Discussion Papers, Federal Reserve Bank of Cleveland, November 2004.

[24] Goto, Akira and Kasuyuki 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* Vol. 71, No. 4, pp. 555-564.

[25] Griffith, Rachel, Stephen Redding, and John van Reenen (2004): 'Mapping the two faces of R&D: Productivity Growth in a Panel of OECD Industries', *Review of Economics and Statistics* Vol. 86, No. 4, pp. 883-895.

[26] Griffith, Rachel, Ruppert Harrison, and John 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* Vol. 96, No. 5, pp. 1859-1875.

[27] Griliches, Zvi (1979): 'Issues in Assessing the Contribution of Research and Development to Productivity Growth', *The Bell Journal of Economics* Vol. 10, No. 1, pp. 92-116.

[28] Grilliches Zvi (1980): 'Returns to Research and Development Expenditures in the Private Sector', in: Kendrick, John W. and Beatrice N. Vaccara (eds.), *New Developments in Productivity Measurement and Analysis*, University of Chicago Press, pp. 419-461.

[29] Griliches, Zvi (1986): 'Productivity, R&D and Basic Research at the Firm Level in the 1970s', *American Economic Review* 76, pp. 141-154.

[30] Griliches, Zvi (1992): 'The Search for R&D Spillovers', *Scandinavian Journal of Economics* Vol. 94, Supplement, pp. S29-S47.

[31] Griliches, Zvi and Frank R. Lichtenberg (1984): 'Interindustry Technology Flows and Productivity Growth: a Reexamination', *Review of Economics and Statistics* Vol. 66, No. 2, pp. 325-329.

[32] Grossman, Gene and Elhanan Helpman (1991): 'Quality ladders in the theory of growth', *Review of Economic Studies* Vol. 58, No. 1, pp. 43-61.

[33] Guellec, Dominique and Bruno van Pottelsberghe (2004): 'From R&D to Productivity Growth: Do the Institutional Settings and the Source of Funds of R&D Matter?', *Oxford Bulletin of Economics and Statistics* Vol. 66, No. 3, pp. 353-378.

[34] Hall, Bronwyn and Jacques Mairesse (1995): 'Exploring the relationship between R&D and productivity in French manufacturing firms', *Journal of Econometrics* Vol. 65, No. 1, pp. 263-293.

[35] Hall, Bronwyn, Jacques Mairesse and Pierre Mohnen (2009): 'Measuring the Returns to R&D', NBER Working Paper No. 15622, December 2009.

[36] Hall, Bronwyn (1996): 'The Private and Social Returns to Research and Development', in: Smith, Bruce L. and Claude E. Barfield (eds.) *Technology, R&D, and the Economy*, Brookings Institution/American Enterprise Institute.

[37] Hall, Bronwyn (2007): 'Measuring the Returns to R&D: The Depreciation Problem', NBER Working Paper No. 13473, October 2007.

[38] Helmers, Christian, Christian Schulte and Hubert Strauss (2009): 'R&D expenditure and capital in Europe', *EIB Papers* Vol. 14, No. 1.

[39] Holly, Sean, M. Hashem Pesaran and Takashi Yamagata (2010): 'A spatio-temporal model of house prices in the USA', *Journal of Econometrics* Vol. 158, No. 1, pp. 160-173.

[40] Kao, Chihwa (1999): 'Spurious regression and residual-based tests for cointegration in panel data', *Journal of Econometrics* Vol. 65, No. 1, pp. 9-15.

[41] Kao, Chihwa, Min-Hsien Chiang and Bangtian Chen (1999): 'International R&D Spillovers: An Application of Estimation and Inference in Panel Cointegration', *Oxford Bulletin of Economics and Statistics* Vol. 61, No. S1, pp. 691-709.

[42] Kapetanios, George, M. Hashem Pesaran and Takashi Yamagata (2011): 'Panels with Nonstationary Multifactor Error Structures', *Journal of Econometrics* Vol. 160, No. 2, pp. 326-348.

[43] Keller, Wolfgang (1998) 'Are International R&D Spillovers Trade Related? Analyzing Spillovers among Randomly Matched Trade Partners', *European Economic Review* Vol. 42, No. 8, pp. 1469-1481.

[44] Keller, Wolfgang (2002a): 'Trade and the Transmission of Technology', *Journal of Economic Growth* Vol. 7, No. 1, pp. 5-24.

[45] Keller, Wolfgang (2002b): 'Geographic localization of international technology diffusion', *American Economic Review* Vol. 92, No. 1, pp. 120-142.

[46] Keller, Wolfgang (2004): 'International Technology Diffusion', *Journal of Economic Literature* Vol. 42, No. 3, pp. 752-782.

[47] Khan, Mosahid and Kul B. Luintel (2006): 'Sources of Knowledge and Productivity: How robust is the Relationship?', OECD Science, Technology and Industry Working Paper 2006/6.

[48] Lichtenberg, Frank and Bruno van Pottelsberghe (1998): 'International R&D Spillovers: A Comment', *European Economic Review* Vol. 42, No. 8, pp. 1483-1491.

[49] López-Pueyo, Carmen, Jaime Sanaú and Sara Barcenilla Visús (2008): 'International R&D spillovers and manufacturing productivity: A panel data analysis', *Structural Change and Economic Dynamics* Vol. 19, No. 2, pp. 152-172.

[50] Luintel, Kul B. and Mosahid Khan (2004): 'Are International R&D Spillovers Costly for the United States?', *Review of Economics and Statistics* Vol. 86, No. 4, pp. 896-910.

[51] Maddala, G.S. and Shaowen Wu (1999): 'A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test,' *Oxford Bulletin of Economics and Statistics* Vol. 61, No. S1, pp. 631-652.

[52] Mairesse, Jacques and Bronwyn 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 No. 5501, March 1996.

[53] Nickell, Stephen (1981): 'Biases in Dynamic Models with Fixed Effects', *Econometrica* Vol. 49(6), pp. 1417-1426.

[54] Odagiri, Hiroyuki (1985): 'Research Activity, Output Growth and Productivity Increase in Japanese Manufacturing Industries', *Research Policy* Vol. 14, No. 3, pp. 117-130.

[55] OECD (2009): 'Reviewing ANBERD and Improving the Measurement of Industrial R&D', Working Party of National Experts on Science and Technology Indicators, DSTI/EAS/STP/NESTI(2009)15.

[56] Park, Walter G. (1995): 'International R&D Spillovers and OECD Economic Growth', *Economic Inquiry* Vol. 33, No. 4, pp. 571-591.

[57] Pesaran, M. Hashem (2004): 'General Diagnostic Tests for Cross Section Dependence in Panels', Cambridge Working Papers in Economics No. 0435, University of Cambridge, June 2004.

[58] Pesaran, M. Hashem (2006): 'Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure', *Econometrica* Vol. 74, No. 4, pp. 967-1012.

[59] Pesaran, M. Hashem (2007): 'A simple panel unit root test in the presence of cross-section dependence', *Journal of Applied Econometrics* Vol. 22, No. 2, pp. 265-312.

[60] Pesaran, M. Hashem and Ron P. Smith (1995): ‘Estimating long-run relationships from dynamic heterogeneous panels’, *Journal of Econometrics* Vol.68, No. 1, pp. 79-113.

[61] Pesaran, M. Hashem and Elisa Tosetti (2011): ‘Large Panels with Spatial Correlations and Common Factors’, *Journal of Econometrics* Vol. 161, No. 2, pp. 182-202.

[62] Romer, Paul (1990): ‘Endogenous Technological Change’, *Journal of Political Economy* Vol. 98, No. 5, pp. 71-101.

[63] Roodman, David (2009): ‘A note on the theme of too many instruments’, *Oxford Bulletin of Economics and Statistics* Vol. 71, No. 1, pp. 135-158.

[64] Schankerman, Mark (1981): ‘The Effects of Double-Counting and Expensing on the Measured Returns to R&D’, *Review of Economics and Statistics* Vol. 63, No. 3, pp. 454-458.

[65] Scherer, Frederic M. (1982): ‘Interindustry technology flows and productivity growth’, *Review of Economics and Statistics* Vol. 64, No. 4, pp. 627-634.

[66] Van Pottelsberghe, Bruno and Frank Lichtenberg (2001): ‘Does Foreign Direct Investment Transfer Technology Across Borders?’, *Review of Economics and Statistics* Vol. 83, No. 3, pp. 490-497.

[67] Wakelin, Katharine (2001): ‘Productivity growth and R&D expenditure in UK manufacturing firms’, *Research Policy* Vol. 30, No. 7, pp. 1079-1090.

[68] Wallis, Gavin and Alex Turvey (2009): ‘Volume of capital services: estimates from 1950 to 2007’, *Economic & Labour Market Review* Vol. 3, No. 1, pp. 53-61.

[69] Wolff, Edward N. and M. Ishaq Nadiri (1993): ‘Spillover Effects, Linkage Structure, and Research and Development’, *Structural Change and Economic Dynamics* Vol. 4, No. 2, pp. 315-331.

Endnotes

[1]In this paper we focus entirely on R&D conducted by the business enterprise sector.

[2]We use the terms productivity and TFP interchangeably throughout this paper to describe the residual of a production function.

[3]R&D is treated as an intermediate input for firms and as current consumption for governments and non-profit organizations (Edworthy and Wallis, 2007). Following the changes to the System of National Accounts in 2008 it is now recommended to treat existing and past R&D as an asset which is capitalized through 'satellite accounting'. The principal motivation for treating R&D expenditure as investment in National Accounting is to compute its contribution to growth in real GDP.

[4]A comprehensive overview of earlier work can be found in Cameron (1996), while Hall et al. (2009) cover more recent studies.

[5]Alternatively, returns to R&D can be obtained directly from using R&D expenditure albeit under the restrictive assumption that knowledge does not depreciate (Hall et al., 2009).

[6]A spatial econometric approach would capture spillovers by imposing a specific structure on the 'spatial' association between countries and/or industries by means of a spatial weight matrix, where the relevant 'space' can be defined in many ways such as geographical, technological, or input-output-based. However, the specification of the spatial weight matrix, which simply produces weighted averages of the R&D variable, remains essentially arbitrary.

[7]Crepon et al. (1998) stress the point that not innovation input (R&D) is supposed to affect productivity, but innovation output. In common with a large number of empirical studies, they use patents as a measure for knowledge output. This however seems too narrow a measure, since knowledge output can also assume many other forms (new products, capital goods, or improved managerial practices). Since R&D is underlying these different innovative outputs, it may be a better and more comprehensive measure of innovation than restricting the analysis to patented innovations.

[8]We are grateful to an anonymous referee for highlighting the problems introduced by double counting and expensing.

[9]Motivation for technology heterogeneity of this type can be taken from the 'new growth' literature (e.g. Aziarides and Drazen, 1990; Banerjee and Newman, 1993) which has resulted in a limited empirical literature (see Eberhardt and Teal, 2011).

[10]This phenomenon must not be confounded with targeted knowledge transfer, e.g. technology transfer within (international) business groups.

[11]Our literature review in Table A-I of the Appendix contains more details and additional studies.

[12]The literature on productivity analysis at the micro-level refers to this as 'transmission bias', which arises from firms' reaction to unobservable productivity realisations when making input choices. Solutions to this problem are then sought via instrumentation of one form or another (for a recent survey of the literature see Eberhardt and Helmers, 2010).

[13]We exclude industry SIC 23 (*Coke, refined petroleum products and nuclear fuel*) for which several countries do not report data.

[14]The selection of countries is determined by data availability. Note that we use data for Germany only after its reunification in 1990.

[15]Interestingly, testing the residuals from a *pooled* AR(2) regression for each of the variables cannot reject cross-section independence for value-added, labour and capital stock, whereas these tests do reject for residuals from *country-specific* AR(2) regressions — the R&D stock

variable however displays substantial cross-section dependence throughout all of these testing procedures possibly indicating the presence of R&D spillovers and other cross-section dependencies.

[16]The empirical analysis was carried out in Stata 10 and we employed a number of user-written Stata routines: `multipurt`, `xtcd` and `xtmg` by Markus Eberhardt; `pescadf` by Piotr Lewandowski; `xtfisher` by Scott Merryman; `abar` and `xtabond2` by David Roodman; `md_ar1` by Måns Söderbom. Routines are available through SSC or the authors' personal webpages.

[17]Our main focus is on the most flexible setup where each country-industry can have a different set of technology parameters. Results for alternatives (country- or industry-level homogeneity) are available on request.

[18]We also considered the Arellano and Bond (1991) estimator, which commonly performs poorly when data are highly persistent (results available on request).

[19]If this assumption is violated no instrument (internal or external) exists which can satisfy both the conditions of validity and informativeness (Pesaran and Smith, 1995).

[20]Note that in standard fashion in all but the POLS models we account for time-invariant unobservables (fixed effects) using dummy variables or model transformations (e.g. first differencing).

[21]These estimators are further remarkably robust to structural breaks, lack of cointegration and certain serial correlation.

[22]Note that in the case of multiple covariates we construct cross-section averages for each in turn: $\bar{x}_{1t}, \bar{x}_{2t}, \ldots, \bar{x}_{kt}$.

[23]We use the arithmetic mean, but it is notable that weighted means can be applied here, provided they submit to certain (granularity) conditions (Pesaran, 2006).

[24]As is so often the case in long-T panels, the AB (available on request) and BB results are very fragile and are dependent on the lag structure chosen for instrumentation. In the AB model we use lagged levels of y_{it}, l_{it}, k_{it} and r_{it} dated $t-3$ and earlier as instruments in the first difference equation, collapsing the instrument matrix to avoid overfitting bias (Bowsher, 2002). We then applied the same strategy in the BB model (in addition we employ lagged differences as instruments in the levels equation) but tested a considerable number of alternatives. In none of the latter did we obtain a coefficient of R&D stock in excess of .05, all of which were statistically insignificant.

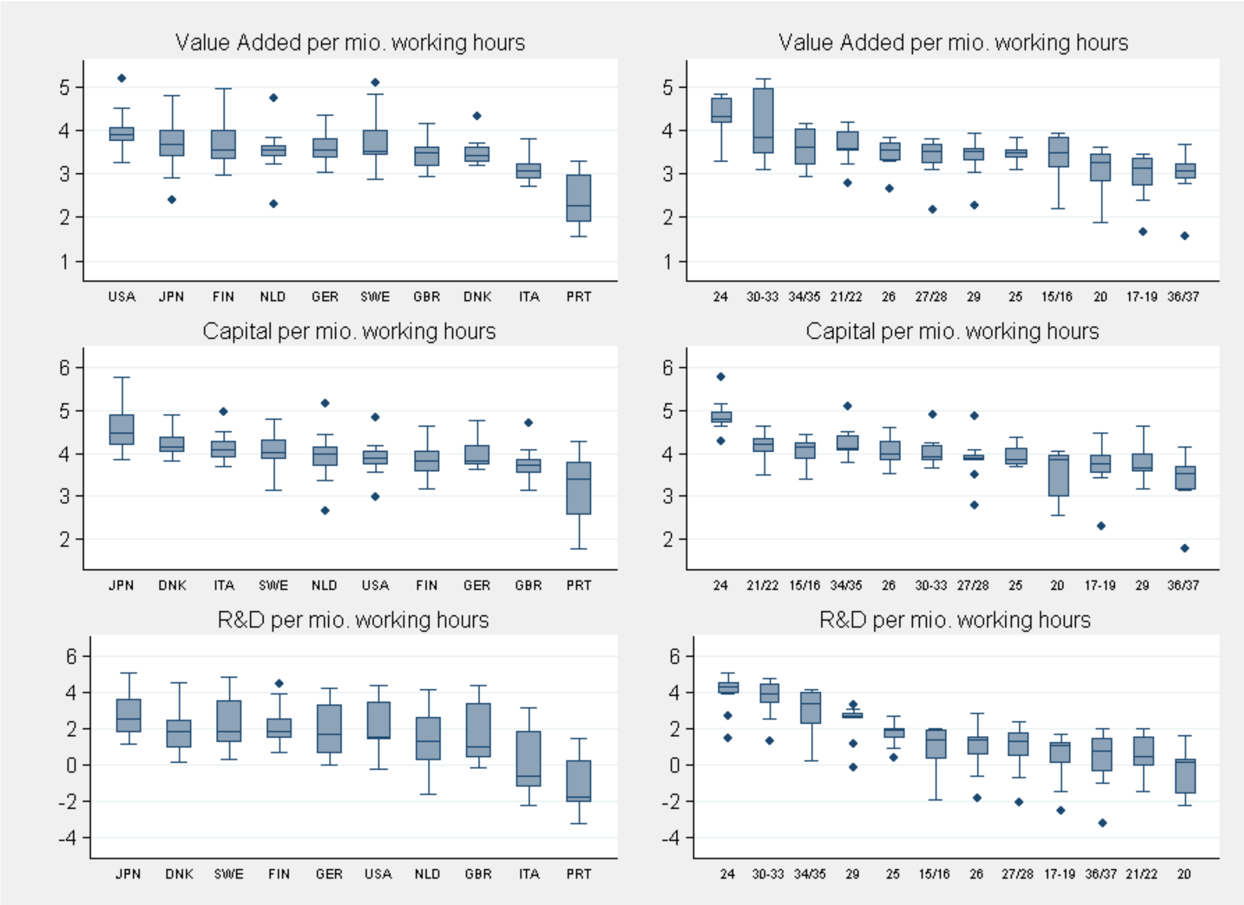
[25]We also conducted these MG-type regressions (static and dynamic) using (outlier-robust) weighted averages instead of the unweighted averages reported in Tables 7 and 8 — findings are qualitatively identical and confirm that our results are not driven by outliers.

[26]The results reported are based on long-run coefficients calculated from the average coefficients in the ARDL model. When we calculate long-run coefficients in each industrial sector and average these, the results are qualitatively the same.

[27]The practical problem consists in splitting 'knowledge spillovers' from 'common shocks' and other cross-section dependencies. For instance, the use of the CCE estimators in a dedicated spatial econometric model fails to recognise that the cross-section averages included in the specification already account for both common shocks and spillovers. It is however anticipated that theoretical developments in this field of research will offer appropriate alternative methods in the near future.

FIGURES AND TABLES

Figure 1: Labour-Deflated Input Variation across Countries & Industries



Notes: The data is transformed into mio. Euros per mio. working hours (in logs) and plotted in order of median value. The left column plots variation by country, the right column by SIC 2-digit industry. All data presented in this graph are for 2005. Dots indicate outliers.

Table 1: Sample makeup: Countries

	Country	Obs	Share	Coverage
DNK	Denmark	312	12%	1980-2005
FIN	Finland	312	12%	1980-2005
GBR	Great Britain	308	12%	1980-2005
GER	Germany	180	7%	1991-2005
ITA	Italy	312	12%	1980-2005
JPN	Japan	312	12%	1980-2005
NLD	Netherlands	312	11%	1980-2005
PRT	Portugal	121	5%	1995-2005
SWE	Sweden	156	6%	1993-2005
USA	United States	312	12%	1980-2005
Total		2,637	100%	

Table 2: Sample makeup: Industries

SIC	Description: Manufacture of	Obs
15, 16	Food, beverages, tobacco	221
17, 18, 19	Textiles, textile products, leather and leather products	221
20	Wood and products of wood and cork	219
21, 22	Pulp, paper, paper products, printing and publishing	219
24	Chemicals and chemical products	221
25	Rubber and plastic products	210
26	Other non-metallic mineral products	221
27, 28	Basic metals and fabricated metal products	221
29	Machinery and equipment n.e.c.	221
30, 31, 32, 33	Electrical and optical equipment	221
34, 35	Transport equipment	221
36, 37	Manufacturing n.e.c.	221
Total		2,637

Notes: Industrial sector SIC 23 (*coke, refined petroleum products and nuclear fuels*) is excluded.

Table 3: Summary statistics

	Mean	Median	Std. Dev.	Min	Max
Levels					
Value-Added (mio. Euro)	27,805	7,992	52,554	290	782,206
Labour (mio. hours worked)	917	393	1,219	15	6,612
Physical Capital (mio. Euro)	40,462	14,535	64340	242	459,870
R&D Capital (mio. Euro)	13,184	846	39,998	0.4	328,954
Logarithms					
ln Value-Added (ln Y_{it})	8.987	8.986	1.683	5.668	13.570
ln Labour (ln L_{it})	5.821	5.974	1.554	2.684	8.797
ln Physical Capital (ln K_{it})	9.431	9.584	1.669	5.487	13.039
ln R&D Capital (ln R_{it})	6.881	6.741	2.505	-0.937	12.704
Growth Rates					
Δ ln Value-Added	0.018	0.015	0.072	-0.412	1.081
Δ ln Labour	-0.015	-0.013	0.044	-0.269	0.185
Δ ln Physical Capital	0.020	0.017	0.031	-0.134	0.213
Δ ln R&D Capital	0.037	0.031	0.064	-0.125	0.790

Notes: These descriptive statistics refer to the sample for $N = 119$ country-industries (from OECD 10 countries), which in levels contains $n = 2,637$ observations, average $T = 22.2$ (range 1980-2005).

Table 4: Overview of Empirical Approach

		Impact of Unobservables: common	heterogeneous
Technology Parameters:	common	*Static*	
		POLS, 2FE, FD	CCEP
		Dynamic	
		POLS, 2FE, BB	CCEP
	heterogeneous	*Static*	
		CDMG	MG, CMG
		Dynamic	
		CDMG	MG, CMG

Notes: POLS – Pooled OLS (with year fixed effects), 2FE – 2-way Fixed Effects, FD – First-Difference OLS, BB – Blundell and Bond (1998), CCEP – Pooled Pesaran (2006) Common Correlated Effects (CCE), MG – Pesaran and Smith (1995) Mean Group, CDMG – Cross-Section Demeaned Mean Group, CMG – Pesaran (2006) CCE Mean Group version.

Table 5: Pooled Production Functions (static)

	POLS [1]	**2FE** [2]	**FD** [3]	**CCEP** [4]	**CCEP** [5]
ln L_{it}	0.464	0.608	0.634	0.563	0.582
	[40.72]**	[18.41]**	[18.01]**	[19.01]**	[19.01]**
ln K_{it}	0.465	0.487	0.274	0.295	0.203
	[37.59]**	[10.60]**	[3.66]**	[7.08]**	[4.45]**
ln R_{it}	0.096	0.063	0.050	0.083	0.064
	[22.80]**	[4.42]**	[1.88]	[4.33]**	[3.30]**
Year dummies	Included	Implicit	Included		Included
CRS	0.00	0.34	0.65	0.15	0.00
AB Test AR(1)	0.00	0.00	0.00	0.00	0.00
AB Test AR(2)	0.00	0.00	0.02	0.18	0.19
CD Test	0.12	0.14	0.21	0.01	0.06
Order of integration	I(1)	I(1)	I(0)	I(0)	I(0)
RMSE	0.278	0.163	0.064	0.059	0.059
Observations	2,637	2,637	2,518	2,637	2,637
Country-industries	119	119	119	119	119

Notes: POLS — Pooled OLS, 2FE — Two-way Fixed Effects, FD — OLS with variables in First Differences, CCEP — Pooled Pesaran (2006) estimator. Absolute t-statistics in brackets, constructed from White heteroskedasticity-robust standard errors. *, ** indicate significance at the 5% and 1% level respectively.
Diagnostics: CRS: Wald test for H_0 of constant returns to scale (labour, physical capital *and* R&D capital; p-values reported). AB Test: Arellano and Bond (1992) test for H_0 of no residual serial correlation (p-values). CD Test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals (p-values). The order of integration of the residuals is determined using the Pesaran (2007) CIPS Test (full results available on request): I(0) – stationary, I(1) – nonstationary, I(1)/I(0) – ambiguous result.

Table 6: Pooled Production Functions (dynamic)

	POLS [1]	**2FE** [2]	**BB** [3]	**CCEP** [4]	**CCEP** [5]
PANEL A: LONG-RUN COEFFICIENTS (UNRESTRICTED MODELS)					
Labour	0.338		-0.524	0.415	0.418
	[2.48]*		[0.80]	[5.90]**	[5.55]**
Capital	0.173		0.894	0.404	0.370
	[0.86]		[1.86]	[4.12]**	[3.46]**
R&D stock	0.462		0.309	0.037	0.032
	[2.77]**		[1.47]	[0.95]	[0.81]
PANEL B: LONG-RUN COEFFICIENTS (RESTRICTED MODELS)					
Labour		0.657			
		[17.24]**			
Capital		0.086			
		[1.35]			
R&D stock		0.024			
		[0.96]			
Year dummies	included	implicit	included		included
COMFAC	0.00	0.73	0.03	0.01	0.02
CRS	0.60	0.56	0.36	0.14	0.10
CD Test	0.13	0.10	0.02	0.61	0.63
Sargan			0.00		
Order of integration	I(1)/I(0)	I(0)	I(1)/I(0)	I(0)	I(0)
RMSE	0.060	0.055	0.053	0.035	0.035
Observations	2,518	2,518	2,518	2,518	2,518
Country-industries	119	119	119	119	119

Notes: BB — Blundell-Bond (1998) System GMM estimator. See also Table 5 for details of tests and other estimators. Absolute t-statistics in brackets, constructed from White heteroskedasticity-robust standard errors. *, ** indicate significance at the 5% and 1% level respectively.
Diagnostics: COMFAC: p-values for H_0 of valid common factor restrictions. All tests (except CRS) are based on the unrestricted ARDL regression results (available on request). Panel A reports unrestricted long-run coefficients, for which standard errors were computed using the Delta method. Panel B imposes the common factor restrictions ex post (provided the COMFAC test indicates the restriction is valid) based on a minimum distance procedure.

Table 7: Heterogeneous production functions (static)

	MG	CDMG	CMG	CMG
	[1]	[2]	[3]	[4]
ln L_{it}	0.568	0.557	0.599	0.698
	[6.57]**	[7.63]**	[9.00]**	[8.24]**
ln K_{it}	0.117	0.445	0.244	0.149
	[0.96]	[5.01]**	[1.70]	[1.00]
ln R_{it}	-0.058	0.089	0.035	-0.050
	[0.73]	[2.12]*	[0.44]	[0.60]
trends	included			included
CRS	0.00	0.09	0.47	0.28
Ljung-Box AR	0.00	0.00	1.00	1.00
Order of integration	I(1)/I(0)	I(1)/I(0)	I(0)	I(1)/I(0)
CD Test	0.00	0.05	0.51	0.35
RMSE	0.051	0.068	0.037	0.035
Observations	2,637	2,637	2,637	2,637
Country-industries	119	119	119	119

Notes: Estimators: MG — Mean Group, CDMG — Cross-sectionally demeand MG, CMG — Pesaran (2006) Common Correlated Effects MG. Absolute t-statistics in brackets, following Pesaran and Smith (1995). *, ** indicate significance at the 5% and 1% level respectively. All averages reported are unweighted means.
Diagnostics: Ljung-Box AR reports the p-values of Fisher statistics constructed from country-industry specific Portmanteau (Q) tests of the residual series for the H_0 of independently distributed residuals/no serial correlation (joint test for up to 3 lags).

Table 8: Heterogeneous production functions (dynamic)

	MG [1]	CDMG [2]	CMG [3]	CMG [4]
LONG-RUN COEFFICIENTS (RESTRICTED MODELS)				
Labour	0.703	0.567	0.642	0.678
	[6.15]**	[10.01]**	[9.39]**	[9.43]**
Capital	0.277	0.245	0.276	0.172
	[1.87]	[3.37]**	[1.70]	[1.09]
R&D stock	-0.107	0.139	-0.084	-0.088
	[0.95]	[3.95]**	[0.94]	[0.96]
trends	included			included
COMFAC	0.72	0.48	0.96	0.85
CRS	0.09	0.19	0.00	0.00
Order of integration	I(1)/I(0)	I(1)/I(0)	I(0)	I(0)
CD Test	0.00	0.07	0.08	0.06
RMSE	0.035	0.038	0.022	0.021
Observations	2,518	2,518	2,096	2,096
Country-industries	119	119	84	84

Notes: See Tables 6 and 7 for details. Absolute t-statistics in brackets, following Pesaran and Smith (1995). *, ** indicate significance at the 5% and 1% level respectively. All averages reported are unweighted means. The common factor restrictions cannot be rejected in any of the four models, we therefore only report the restricted model results (ARDL results available on request). We dropped data from SWE, GER and PRT for the CMG models due to the dimensionality problem (MG and CDMG estimates for smaller sample match those presented).

APPENDIX

A Literature Overview

Table A-I provides an overview of the literature on returns to own R&D and R&D-related spillovers based on the recent survey article on the measurement of returns to R&D by Hall et al. (2009). The selection reported here, however, is much smaller because we focus on articles that match our own approach as closely as possible. The selection criteria are as follows: (i) country or country-industry as the unit of observation (i.e. no firm-level studies); (ii) production function setup (as opposed to cost functions); (iii) samples of developed countries; (iv) studies featuring an explicit own-R&D variable for the business sector (as opposed to specifications using total R&D or confounding own and others' R&D); and (v) published after 1980. Our selection includes some 30 articles of which the 19 studies deemed most relevant are summarized in Table A-I. The full table including all studies as well as further information on the specifics of the underlying production function, the estimators, alternative specifications, additional variables and results is available from the authors upon request.

In Table A-I, we make a distinction between country-level and industry-level studies. The latter are further divided into those analysing *inter-industry* spillovers, whereby the sample may contain only one country (central part of the table), and articles with a particular — albeit not exclusive — interest in *international spillovers within the same industry*. As shown in the table, most studies first derive TFP indices from a standard growth accounting framework and regress this TFP index on own R&D, the R&D spillover variable and sometimes a set of control variables and dummies. When the R&D variables come as R&D capital stocks, results are to be interpreted as elasticities and reported in column 'Elasticity'. When R&D variables are intensities (R&D/value-added), results are interpreted as gross returns (assuming zero depreciation, see Hall et al., 2009) and are reported in column 'Gross return'. Column 'Modeling of Spillovers' summarises the different weighting schemes employed to aggregate R&D in other industries/countries into an R&D spillover variable. Studies can be broadly divided into employing spillover measures aimed at uncovering rent spillovers or knowledge spillovers. In the former case, input-output relations between industries and trade relations between countries are used. In the latter case, preference is given to patent flows across industries and countries and indicators of countries' similarity in research field composition among others. The spillover coefficients are reported in the final column.

Table A-I: Literature Review

Author(s) (year)	Sample	Years	Dependent variable	Modeling of spillovers	Estimated R&D coefficient: Elasticity	Estimated R&D coefficient: Gross return	Spillover coefficient
I. Cross-country studies (aggregate economy level)							
Coe & Helpman (1995)	22	1971-90	log(TFP)	FRDS (import-share-weighted) interacted with import/GDP	0.23 (G7), 0.08 (others)		0.29
Keller (1998)	22	1971-90	TFP	FRDS (random weights) interacted with import/GDP	0.13 (G7), 0.035 (others)		0.05 (G7 to others)
Kao, Chiang & Chen (1999)	22	1971-90	TFP	FRDS (import-share-weighted) interacted with import/GDP	0.21 (G7), 0.08 (others)		0.26
Van Pottelsberghe & Lichtenberg (2001)	13	1971-90	log(TFP)	FRDS weighted by imports, outw. or inw. FDI	0.05 0.06 0.08		0.067 (import weights) 0.039 (outw. FDI weights) [0.006] (inw. FDI weights)
Luintel & Khan (2004)	10	1965-99	log(TFP)	FRDS (import-share-weighted) interacted with import/GDP	0.29 (avg LR)		0.12 (avg LR)
Guellec & van Pottelsberghe (2004)	16	1980-98	TFP	FRDS technological-proximity weighted (technology-distance)	0.13		0.46
Coe, Helpman & Hoffmaister (2009)	24	1971-2004	log(TFP)	FRDS (import-share-weighted) interacted with import/GDP	0.20 (G7), 0.13 (others)		0.065
II.1. Industry studies: Inter-industry spillovers (one or more countries)							
Scherer (1982)	US, 87 m	1959-78	TFP	Product R&D of other industries, patent-flow weighted		1964-69: [0.13] 1973-78: 0.29	1964-69: 0.64 1973-78: 0.74
Griliches & Lichtenberg (1984)	US, 193 m	1959-78	Avg TFP growth of adjacent half-decades	Product R&D of other industries, patent-flow weighted		1959/63-1964/68: 0.29 1964/68-1969/73: 0.11 1969/73-1974/78: 0.31	1959/63-1964/68: 0.51 1964/68-1969/73: 0.90 1969/73-1974/78: 0.50
Odagiri (1985)	Japan, 15 m	1960-77	TFP growth	Sum of R&D expenditures of other industries, weighted by industry's share in total sales		1.57 to 3.15	-6.06 to 7.34
Goto & Suzuki (1989)	Japan, 50 m	1978-83	Avg annual TFP growth	Inter-industry transactions		0.25	0.80
Wolff & Nadiri (1993)	US, 50 i	1947/58/63/ 67/72/77	TFP growth rate	R&D in other industries i, weighted by 1) share of deliveries to ind. j in j's gross output 2) contributions to change of j's capital stock		1) 0.17 2) 0.21	1) [0.08] 2) 0.09
Griffith, Redding & Van Reenen (2004)	12 c, 12 i	1974-90	TFP growth rate	1) Industry's TFP gap from frontier 2) Own R&D intensity interacted with TFP gap		1) 0.67 2) 0.50	1) 0.07 2) 0.60
Cameron, Proudman & Redding (2005)	UK, 14 m	1970-92	TFP growth rate	1) Industry's TFP gap from frontier 2) Own R&D intensity interacted with TFP gap		1) 0.70 2) 0.64	1) 0.10 2) [0.66]
II.2. Industry studies: Cross-country intra-industry spillovers							
Braconnier & Sjöholm (1998)	6 c, 9 m	1979-91	TFP growth rate	Other industries' R&D weighted by 1) dom. intermediates 2) imported intermediates 3) Foreign R&D in same industry (FDI-weighted)	0.03		1)[-0.53] dom. inter-ind. 2) [1.09] int. inter-ind 3) 0.0006 int. intra-ind
Keller (2002a)	8 c, 13 m	1970-91	log(TFP index)	Other country-industry pairs' R&D stocks: i) dom. inter-ind., ii) int. intra-ind., iii) int. inter-ind.; Weights are based on 1) I/O relations/intermediate imports 2) Source & user industry for Canadian patents	1) 0.61 2) 0.15		1) i) 0.57; ii) 0.09; iii) 0.29 2) i) [0.39]; ii) 0.22; iii) [0.19]
Keller (2002b)	14 c, 12 m	1970-95	log(own TFP/ weighted avg foreign TFP)	Sum of cumulative R&D expenditure weighted by bilateral distance	0.078		0.84
Baldwin, Braconier & Forslid (2005)	9 c, 7 m	1979-91	(VA/employee) growth	Int. intra-industry spillover 1) FDI spillover 2) Marshall-Arrow-Romer spillover	0.001		1) 25.5 2) 0.63
López-Pueyo, Sanaú & Barcenilla Visús (2008)	6 c, 10 m	1979-2000	Level of TFP	Weighted sum of 3 R&D capital stocks: (i) same foreign ind; (ii) other foreign ind.	0.143		i) 0.107; ii) 0.213

Notes: Statistically insignificant estimates are reported in square brackets; FRDS: Foreign R&D stock; c: country; m: manufacturing; i: industry; I/O: input/output; LR: long-run; dom.: domestic; int: international; FDI: foreign direct investment; avg: average (unweighted unless indicated). Full literature review with additional information available on request.

B Variable construction

B-1 Output — Value-added

We use value-added as a measure of industry output in order to achieve comparability with the existing literature and because value-added is more closely related to profitability than sales. EU KLEMS reports both gross output and intermediate inputs in current prices. We therefore construct double-deflated value-added by subtracting real inputs from real output. This practice is preferable over using single-deflated value-added, i.e., deflated nominal value-added, as a measure for output, since it avoids the situation where differential price movements across countries generate the false impression of productivity changes. EU KLEMS also provides the necessary industry-level deflators which is a distinct advantage of this data as for some industries, expectations of price changes would likely be different to the general level of inflation. This is an important issue because if inadequate deflators are used, industry output may appear to grow slower. Since this is most likely in industries that are R&D-intensive, the contribution of R&D to output growth would be underestimated (Hall, 1996).[28]

To account for the 'expensing bias' discussed in Section 2 of the main text, we adjust intermediate inputs for R&D-related expenses. We use OECD data to construct the share of intermediate R&D inputs in total R&D expenditure to adjust the conventional measure of intermediate inputs. We then use this adjusted intermediate input measure to construct our double-deflated measure of value-added. In an alternative specification we include the measure for R&D intensity directly in our regression model. The sample for the Schankerman-adjustments (we only discuss the coverage for the final regression sample which is diminished primarily by the lack of industry-level data on R&D workers) covers 97 country-industries in 9 countries (SWE is missing, USA only has 5 observations in one industry) between 1987 and 2005 (with 1988, 1992, 1994, 1996 further missing) and has a total of 725 observations, thus less than 30% of the full sample analysed in the main section of the paper.

B-2 Labour input

As a measure of labour input, EU KLEMS provides the total number of hours worked by persons engaged. The availability of such information is an advantage of EU KLEMS over other datasets as usually the number of full-time equivalent employees has to serve as a proxy for labour input, possibly aggravating the problem of measurement error (see for example Hall and Mairesse, 1996; Wakelin, 2001).

In order to correct for 'double counting' of R&D in our measure of labour input as suggested by Schankerman (1981), we construct the ratio of R&D labour input and traditional labour. The data come from EUKLEMS and the OECD.

B-3 Capital input

Ideally, a measure of current capital services instead of capital stocks, i.e., a flow measure instead of a stock measure, should be used in productivity analysis (Jorgenson and Griliches, 1967).[29] The EU KLEMS dataset provides such a measure for capital services in index form. However, since we do not have any data on R&D capital services, we prefer to use physical capital stocks as a proxy for capital services.[30] This is acceptable under the assumption that the quantity of an asset held by an industry is proportional to the quantity of the corresponding service obtained from that asset. For this to be the case, the aggregate of an industry's capital holdings should represent an average over the various different vintages and age groups of the capital employed within the industry. That this assumption may approximately hold in practice is supported by empirical work, for example, by Wallis and Turvey (2009) for the UK.

Capital input is measured as total tangible assets by book value recorded annually. EU KLEMS provides several measures for tangible assets including total tangible assets, gross fixed capital formation (GFCF), ICT assets, and non-ICT assets. We use total tangible assets and deflate them using a industry-level producer price index.

We create a measure for the ratio of physical capital devoted to R&D and total physical capital to correct for 'double-counting' of R&D (Schankerman, 1981). The measure is constructed as the share of R&D spending on capital in total R&D spending where the data come from the OECD.

B-4 R&D expenditure and stocks

We use R&D stocks in our analysis. It is well known that R&D takes time to translate into innovation and it is therefore the ensemble of past and current R&D expenditures that should matter for productivity rather than merely current expenditure. At the same time past knowledge also depreciates, hence, simply specifying lagged *levels* of R&D expenditure to account for the dynamic nature of R&D may be misleading. The combination of knowledge accumulation and depreciation is also the underlying rationale for Equation (2) in the Griliches knowledge production framework: the notion that more recent vintages of R&D investment matter more for the knowledge stock than older ones is captured by the log polynomial specification.

EU KLEMS provides R&D stocks for 19 countries for the period 1980-2003. However, the overlap with the available tangible capital stock data is not perfect leaving us with 9 countries for which both R&D stocks and physical capital data are available. In order to increase the number of countries in the sample, we constructed R&D capital stocks for Portugal for which R&D data is readily available. These R&D stocks were computed using the OECD Analytical Business Enterprise Research & Development (ANBERD) data (update May 2009) which only accounts for business enterprise R&D.[31] EU KLEMS also uses ANBERD to construct R&D stocks and we followed their methodology for Portugal applying the perpetual inventory method (PIM):

$$R_{it} = (1-\delta)R_{it-1} + \text{R\&D}_{it} \tag{19}$$

where R&D denotes real R&D flows and R the corresponding stock. In order to implement equation (19), δ has to be determined. In line with EU KLEMS, we assume a depreciation rate of 12% (Hall and Mairesse, 1995; Hall, 2007). The depreciation rate is assumed to be the same across industries and constant over time: as noted by Hall and Mairesse (1995), the actual rate chosen seems to be of little relevance for estimation. The reason is the same that also justifies the use of the following formula to compute the initial capital stock

$$\begin{aligned} R_{i1} &= \text{R\&D}_{i0} + (1-\delta)\ \text{R\&D}_{i-1} + (1-\delta)^2\ \text{R\&D}_{i-2} + \ldots \\ &= \sum_{t=0}^{\infty}(1-\delta)^t\ \text{R\&D}_{i-s} = \text{R\&D}_{i0}\sum_{t=0}^{\infty}\left[\frac{1-\delta}{1+g_i}\right]^t = \frac{\text{R\&D}_{i0}}{\delta+g_i} \end{aligned} \tag{20}$$

where g_i denotes the industry-specific growth rate of R&D capital stock. Contrary to other authors, such as Hall and Mairesse (1995), we do not assume a value for g_i but compute it using the first seven years for which R&D expenditure is observed. As long as the growth rate and the depreciation rate do not change dramatically within industries over time, they will be captured by industry-specific effects in any regression. Hence, the elasticity of output with respect to R does not depend on the choice of δ.

In addition to constructing R&D capital stocks for Portugal, we extended the R&D stocks computed by EU KLEMS for all other countries to cover 2004 and 2005 as well, using ANBERD data and PIM described above. We used GDP deflators as proxies for R&D-specific deflators to obtain real R&D expenditures prior to computing the stock variables. We acknowledge a potential measurement problem arising from this choice (see Edworthy and Wallis, 2007) but at present no viable alternative data are available.

Despite efforts undertaken by the OECD to produce internationally comparable R&D data, important differences across countries in their attribution of R&D across industries remain, including data collection, changes in classification and annual data coverage (OECD, 2009). For our data, the problem in international comparability arises from the fact that countries do not report R&D data uniformly by product field but some rather by main activity. Countries also differ in their treatment of R&D conducted in the 'R&D services' industrial sector ISIC 73. Our set of countries contains countries that follow either the product field or main activity approach: Denmark, Germany, Italy, Japan, Netherlands, Portugal and the US follow the main activity approach, whereas Finland, Sweden, and the UK follow the product field approach. This difference in the allocation of R&D spending across industries still contaminates cross-country comparability of R&D expenditures and stocks.[32]

TECHNICAL APPENDIX

TA-1 Variable Properties

Table TA-1: Time-Series Properties

PANEL A: VARIABLES IN LEVELS

Maddala and Wu (1999) Fisher Test

	Constant					Constant and Trend			
lags	$\ln Y_{it}$	$\ln L_{it}$	$\ln K_{it}$	$\ln R_{it}$	lags	$\ln Y_{it}$	$\ln L_{it}$	$\ln K_{it}$	$\ln R_{it}$
0	377.10 (.00)	195.89 (.98)	475.55 (.00)	821.56 (.00)	0	237.33 (.50)	165.30 (1.00)	218.26 (.82)	113.11 (1.00)
1	387.37 (.00)	318.94 (.00)	353.65 (.00)	376.22 (.00)	1	448.85 (.00)	405.17 (.00)	381.98 (.00)	585.53 (.00)
2	329.96 (.00)	184.69 (1.00)	277.02 (.04)	373.42 (.00)	2	337.86 (.00)	233.73 (.57)	254.39 (.22)	210.84 (.90)
3	292.94 (.01)	211.53 (.89)	329.64 (.00)	361.32 (.00)	3	272.38 (.06)	280.36 (.03)	481.75 (.00)	429.02 (.00)

Pesaran (2007) CIPS Test

	Constant					Constant and Trend			
lags	$\ln Y_{it}$	$\ln L_{it}$	$\ln K_{it}$	$\ln R_{it}$	lags	$\ln Y_{it}$	$\ln L_{it}$	$\ln K_{it}$	$\ln R_{it}$
0	2.33 (0.99)	3.46 (1.00)	8.01 (1.00)	9.45 (1.00)	0	1.11 (0.87)	3.45 (1.00)	8.01 (1.00)	10.26 (1.00)
1	2.50 (0.99)	-0.24 (0.41)	8.43 (1.00)	7.13 (1.00)	1	-3.30 (.00)	-1.60 (0.06)	-2.62 (.00)	0.57 (.72)
2	10.36 (1.00)	8.39 (1.00)	10.27 (1.00)	14.58 (1.00)	2	8.47 (1.00)	9.88 (1.00)	6.98 (1.00)	9.52 (1.00)
3	15.22 (1.00)	12.55 (1.00)	11.63 (1.00)	16.51 (1.00)	3	18.73 (1.00)	17.61 (1.00)	14.65 (1.00)	17.53 (1.00)

PANEL B: VARIABLES IN FIRST DIFFERENCE (WITH DRIFT)

	Maddala and Wu (1999) Fisher Test					Pesaran (2007) CIPS Test			
lags	$\ln Y_{it}$	$\ln L_{it}$	$\ln K_{it}$	$\ln R_{it}$	lags	$\ln Y_{it}$	$\ln L_{it}$	$\ln K_{it}$	$\ln R_{it}$
0	1674.68 (.00)	1140.05 (.00)	579.39 (.00)	395.02 (.00)	0	-22.96 (.00)	-16.84 (.00)	-10.22 (.00)	-3.25 (.00)
1	1245.59 (.00)	879.13 (.00)	460.91 (.00)	537.31 (.00)	1	-14.83 (.00)	-11.54 (.00)	-5.27 (.00)	-6.37 (.00)
2	750.23 (.00)	469.34 (.00)	386.03 (.00)	308.01 (.00)	2	-2.19 (.01)	2.16 (.98)	1.66 (.95)	3.79 (1.00)
3	460.06 (.00)	422.29 (.00)	582.17 (.00)	356.46 (.00)	3	12.23 (1.00)	13.42 (1.00)	14.64 (1.00)	10.56 (1.00)

Notes: For the Maddala and Wu (1999) test we report the Fisher statistic and associated p-value, for the Pesaran (2007) test the standardised Z-tbar statistic and its p-value. The null hypothesis for both tests is that all series are nonstationary. Lags indicates the lag augmentation in the Dickey Fuller regression employed. In Panel A we augment the Dickey Fuller regression for variables in levels with a constant or a constant and trend; in Panel B for the variables in first differences we only employ a drift (constant). We used Stata routines `xtfisher` and `pescadf` written by Scott Merryman and Piotr Lewandowski respectively.

Table TA-2: Cross-Section Correlation

PANEL A: LEVELS	$\ln Y_{it}$	$\ln L_{it}$	$\ln K_{it}$	$\ln R_{it}$	PANEL B: FIRST DIFFERENCES	$\Delta\ln Y_{it}$	$\Delta\ln L_{it}$	$\Delta\ln K_{it}$	$\Delta\ln R_{it}$
avg ρ	0.29	0.30	0.55	0.40	avg ρ	0.17	0.17	0.20	0.03
avg $\|\rho\|$	0.59	0.57	0.77	0.78	avg $\|\rho\|$	0.26	0.28	0.34	0.34
CD	110.44	105.45	199.00	149.64	CD	58.78	59.08	68.53	12.50
p-value	0.00	0.00	0.00	0.00	p-value	0.00	0.00	0.00	0.00
PANEL C: POOLED AR(2)	$\ln Y_{it}$	$\ln L_{it}$	$\ln K_{it}$	$\ln R_{it}$	PANEL D: COUNTRY-INDUSTRY AR(2)	$\ln Y_{it}$	$\ln L_{it}$	$\ln K_{it}$	$\ln R_{it}$
avg ρ	0.00	0.00	0.00	0.02	avg ρ	0.13	0.12	0.09	0.02
avg $\|\rho\|$	0.23	0.26	0.24	0.25	avg $\|\rho\|$	0.25	0.25	0.25	0.23
CD	-0.55	-1.42	-1.03	7.05	CD	45.46	42.20	33.78	8.44
p-value	0.58	0.16	0.30	0.00	p-value	0.00	0.00	0.00	0.00

Notes: We present the average and average absolute correlation coefficients across the $N(N-1)$ sets of correlations. CD reports the Pesaran (2004) cross-section dependence statistic, which is distributed $N(0,1)$ under the null of cross-section independence. Panels A and B test the variable series in levels and first differences respectively. In Panel C each of the four variables in levels is entered into a pooled panel regression $z_{it} = \pi_{0,i} + \pi_1 z_{i,t-1} + \pi_2 z_{i,t-2} + \pi_t + \varepsilon_{it}$ where π_t indicates $T-1$ year dummies and $\pi_{0,i}$ N country-industry fixed effects. In Panel D each of the four variables in levels is entered into a time-series regression $z_{it} = \pi_{0,i} + \pi_{1,i} z_{i,t-1} + \pi_{2,i} z_{i,t-2} + \pi_{3,i} t + \varepsilon_{it}$, conducted separately for each country-industry i. The correlations and cross-section dependence statistic in Panels C and D are then based on the residuals from these AR regressions. We used the Stata routine `xtcd` written by Markus Eberhardt.

TA-2 Schankerman (1981) Correction/Augmentation

We carry out variable adjustments to account for excess return bias due to double-counting (DC) and model augmentation to account for expensing bias (EB), following Schankerman (1981). Since our data coverage for s (share of R&D workers in total workforce), δ (share of R&D investment in total investment) and θ (R&D intensity) is relatively limited (we lose over 70% of observations) we are unable to estimate dynamic model specifications and limit our analysis to static models. Furthermore, the augmentations with cross-section averages in the (standard, augmented) CCEP estimators necessitate a sample reduction such that the number of country-industries would drop to a mere 33 ($n = 395$ observations). In addition, these estimators rely on the time-series dimension of the panel to estimate the country-specific coefficients on the cross-section averages and therefore cannot be expected to perform well in the resulting sample setup where T ranges from 8 to 13. We therefore drop these estimators from this robustness exercise.

We present results for a Griliches knowledge production function where input variables are adjusted for *observed values* of s & δ and with *observed* θ included as additional regressor. Alternatively, we use unadjusted input variables and add s, δ, and θ to the standard Griliches knowledge production function to account for the omitted variable bias. We also experimented with adjusting value-added directly by correcting the intermediate input measure for R&D expensing. Results showed very similar patterns to those presented in Table TA-3, Panel A, and are therefore not presented here.

The results for the pooled models where k and l are adjusted and R&D intensity is added as covariate (Table TA-3, Panel A) largely follow the direction of the bias suggested by Schankerman (1981): in all but the POLS models correcting for double-counting raises the coefficient on R&D capital. Further, adjusting for expensing can be seen to have an ambiguous effect across empirical models. When we instead add measures for s and δ to the regression equation with unadjusted k and l (same Table, Panel B) the coefficient on R&D capital hardly moves at all and the tests for the constraints linking the coefficients (see Schankerman (1981) footnote 4) reject in all models.

We further experimented with some data imputations, replacing missing observations with country-industry time-series averages (this yielded $n = 2,292$ observations), but these results proved not to be particularly insightful, following the patterns described in the smaller sample for observed data only.

In conclusion, given all the data constraints experienced we can merely highlight the seemingly limited change in the R&D coefficients once we adjust for expensing and double-counting. From an econometric perspective we believe there are good grounds to suggest that other data properties, first and foremost nonstationarity and cross-section dependence, play an important role in this type of data and that the empirical bias derived by Schankerman (1981) in a cross-section regression of firm-level data may be conflated with a failure to address these more salient macro panel data issues in the present case. For the empirical models which explicitly account for cross-section dependence there is furthermore a theoretical argument that they can address the double-counting and expensing problem (see Section 3.2.2 in the main text).

Table TA-3: Schankerman Correction/Augmentation — Pooled Models

PANEL A: ADJUSTMENT using observed data for s, δ and θ

	POLS			2FE			FDOLS		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
$\ln R_{it}$	0.135	0.132	0.138	0.121	0.163	0.162	-0.005	0.041	0.043
	[14.19]**	[13.26]**	[13.77]**	[1.29]	[1.57]	[1.53]	[0.09]	[0.59]	[0.61]
Correction		DC	DC, EB		DC	DC, EB		DC	DC, EB
Year dummies	included	included	included	included	included	included	included	included	included
Observations	725	725	725	725	725	725	306	306	306
Aveage T	7.3	7.3	7.3	7.3	7.3	7.3	3.1	3.1	3.1

PANEL B: AUGMENTATION using observed data for s, δ and θ

	POLS			2FE			FDOLS		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
$\ln R_{it}$	0.135	0.127	0.131	0.121	0.101	0.099	-0.005	-0.007	-0.005
	[14.19]**	[12.90]**	[13.35]**	[1.29]	[1.09]	[1.07]	[0.09]	[0.11]	[0.08]
Augmentation		DC	DC, EB		DC	DC, EB		DC	DC, EB
Year dummies	included	included	included	included	included	included	included	included	included
Observations	725	725	725	725	725	725	306	306	306
Aveage T	7.5	7.5	7.5	7.5	7.5	7.5	3.2	3.2	3.2
Restricton F-test (s,δ)		5.37	4.45		22.16	22.24		7.13	7.13
p-value		0.005	0.012		0.000	0.000		0.001	0.001

Notes: DC — double-counting (correct variable for/augment model with s, δ), EB — expensing bias (augment model with θ). Year dummies included in all models. Constraint refers to an F-test linking coefficients on s and δ to those on ln L and ln K respectively. See text above for more details on these exercises. $*$, $**$ indicate statistical significance at the 5% and 1% level respectively. $N = 99$ country-industries in all regressions.