DO SPILLOVERS MATTER WHEN ESTIMATING PRIVATE RETURNS TO R&D?

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 twelve 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.

Because 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)

Handbook chapter.

I. 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 other 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, nonexcludability, and nonexhaustability, 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. A vast economic literature attributes an eminent role to R&D in generating productivity gains and long-run growth owing to the generation of spillovers (Romer, 1990; Grossman & 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 (that is, private) returns can be estimated without also accounting for the spillovers it induces.

Considering the importance of the subject, it is not surprising that a substantial number of empirical studies assessing the private and social returns to R&D at the country, regional, industry, and firm levels.[4] A closer look at this literature, which is summarized in table A-1 in the online 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, labor and capital. This also implies that R&D enters the production function in an additively separable way, 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 & Jaumandreu, 2008). In parallel to this approach, a large body of research concentrates on the contribution of spillovers to productivity, imposing a rigid structure on the spillover channels in constructing spillover variables based

Received for publication May 23, 2010. Revision accepted for publication August 10, 2011.

* Eberhardt: University of Nottingham; Helmers: Universidad Carlos III de Madrid; Strauss: European Investment Bank.

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, 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 and session participants for their comments and suggestions. The usual disclaimers apply. M.E. gratefully acknowledges financial support from the U.K. 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.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00272.

[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 nonprofit organizations (Edworthy & 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 that 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); 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).

The Review of Economics and Statistics, May 2013, 95(2): 436–448

So we don't know how spillovers occur.

on somewhat ad hoc assumptions. This practice reflects the general lack of a clear understanding about the precise channels through which (unobservable) spillovers occur.

Research question

spillovers create cross-correlation

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 crosssectional 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 macrolevel when accounting for any unobserved heterogeneities, including local or global spillovers, and common shocks. This means that our results are not based on ad hoc assumptions about the structure of spillovers, and we do not 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 to 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 of 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 or other spillover effects, or both), as well as technology homogeneity across countries 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 sizable; when we account for spillovers of unknown form, which may include factors other 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 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.

This paper suggests that when taking into account common factors, R&D returns are not as big. Equivalently not allowing for spillovers is like an omitted variable bias.

ok, but what about industry, location or any common factor fixed effect?

The remainder of this paper is organized as follows. Section II discusses the theory underlying the Griliches knowledge production function at the heart of the literature. Section III discusses the theory on knowledge spillovers, as well as their empirical measurement. Section IV introduces the data set used for our analysis and provides descriptive statistics. Section V contains a description of our estimation strategy, and section VI presents the empirical results. Section VII concludes.

II. 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 labor L and tangible capital K as well as knowledge capital R:

value added function.

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

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

[6] A spatial econometric approach would capture spillovers by imposing a specific structure on the spatial association between countries 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.

Kinda critiques Lychagin et al (2016) approach of using a weighting matrix approach for the dimensions of the firm.

function: the functional relation between knowledge inputs and knowledge output.[7] Griliches then writes equation (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 lowercase 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 that 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 how the R&D capital stock (R) can be constructed from R&D expenditures ($R\&D$) in the online appendix, section B.4. 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: $\gamma = \frac{\partial Y}{\partial R}\frac{R}{Y}$. 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 labor also contain elements of R&D, which is thus double-counted because R&D workers are included in the total labor force head count 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—the risk premium or supranormal 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 recognize 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 labor 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 & Mairesse, 1995), although the significance of this bias in panel data sets accounting for fixed effects is subject to some debate (Cuneo & Mairesse, 1984; Hall & Mairesse, 1995; Guellec & van Pottelsberghe, 2004). Our strategy to deal with these econometric difficulties will be twofold. First, we show that the unobserved common factor model adopted in our empirics and detailed in section IIIB is theoretically appropriate to tackle the excess returns and expensing biases. Second, we follow Schankerman's (1981) suggestion 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 VI and presented in more detail in the online appendix.

The overall validity of the Griliches knowledge production function approach rests on the assumption of perfectly competitive factor markets, full capacity utilization, and the absence of spillover effects, the last econometrically represented by the cross-sectional independence of the error term 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 sections.

III. 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, primary among which is the specification of a known, additively separable, functional form that 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

main
ensumption

[7] Crepon, Duguet, and Mairesse (1998) stress 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 underlies 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 (Aziarides & Drazen, 1990; Banerjee & Newman, 1993), which has resulted in a limited empirical literature (see Eberhardt & Teal, 2011).

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.

A. *Knowledge Spillovers*

Arrow (1962) pointed out that knowledge is distinct from the traditional production factors labor and physical capital. The distinguishing features are nonexcludability and nonrivalry 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).

B. *Spillovers in the Knowledge Production Function*

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 $i = 1, \ldots, N$ industries within a single country for simplicity of exposition. First, TFP is estimated or computed from value-added and standard factor inputs labor 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:

Lychagin et al estimates like this ↳

$$\mathrm{TFP}_{it} = g\left(R_{it}, \sum_{k=1}^{N} \omega_k R_{kt}\right), \tag{5}$$

where R_{it} denotes the R&D stock of industry i and the second term in parentheses captures spillovers received from all other industries, with ω_k some explicit weights structuring the relative importance of industries.[11] This setup allows 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 & Suzuki, 1989; Keller, 2002a), import weights (Coe & Helpman, 1995; Keller, 1998), inward-outward FDI or shares of foreign affiliates' sales in domestic sales of an industry (van Pottelsberghe & Lichtenberg, 2001; Baldwin, Braconier, & Forslid, 2005), geographic distance (Keller, 2002b), distance to technology frontier as measured by TFP differences (Griffith, Redding, & van Reenan, 2004; Cameron, Proudman, & Redding, 2005; Acemoglu, Aghion, & Zilibotti, 2006), and measures of technological proximity (Conley & Ligon, 2002; Guellec & van Pottelsberghe, 2004).[12]

Equation (5) can be estimated as

$$\mathrm{tfp}_{it} = \psi_i + \overbrace{\gamma r_{it}}^{\text{own R\&D}} + \chi \underbrace{\sum_{k=1}^{N} \omega_k r_{kt}}_{\text{spillovers}} + \varepsilon_{it}, \tag{6}$$

where lowercase 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) that 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 important, 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 that can capture knowledge spillovers of unknown form, together with any other unobserved heterogeneities that may cause cross-sectional correlation, allows for heterogeneous production technology across industries and countries, and is concerned with the appropriate treatment of dynamics and time-series properties more generally.

Unobserved common factor framework. The common factor approach assumes that the error term as well as the

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

[11] By practical convention the weight on own-industry R&D is set to 0 ($\omega_{k=i} = 0$) in this computation.

[12] Our literature review in table A-1 in the online appendix contains more details and additional studies.

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 emphasized 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, Pesaran, & Yamagata, 2010; Chudik, Pesaran, & Tosetti, 2011). This setup has particular appeal for the analysis of returns to own-R&D in a set of interconnected OECD countries that are subject to common shocks that may have a differential impact on individual economies, and where R&D may spill over from one industry or economy to another following a complex, unknown, and nonsymmetric 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 multifactor error structure and the assumed driving force of the input,

common factor

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

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

where e_{it} and ε_{it} are stochastic shocks. The setup assumes that latent processes drive both productivity and 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 equation (7) without accounting for f_t produces biased and inconsistent estimates of $E[\beta_i] = \beta$, as can be shown by simple substitution,

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

you estimate (9) instead of (7) to get β
you get a biased/inconsistent estimate.

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.[13]

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 labor, capital stock, and value-added are mismeasured by the share of R&D workers in total labor (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

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

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

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 (for example, 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):

π_{ki}

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

Since these common factors are correlated with the R&D stock (Schankerman, 1981), 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 VIA).

IV. Data

The data set comprises information on up to twelve manufacturing industries (SIC 15-37 excluding SIC 23)[14] in ten countries (Denmark, Finland, Germany, Italy, Japan, Netherlands, Portugal, Sweden, United Kingdom, and the United States) over a time period of up to 26 years from 1980 to 2005, yielding 2,637 observations (see tables 1 and 2 for details).[15] 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 of sources, including the EU KLEMS data

[13] The literature on productivity analysis at the microlevel refers to this as transmission bias, which arises from firms' reaction to unobservable productivity realizations 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 & Helmers, 2010).

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

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

analysis unit: industry x country
how to compare it to say Lychagin et al. (2016) which

uses firms ?

TABLE 1.—SAMPLE MAKEUP: COUNTRIES

	Country	Observations	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	Observations
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

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

set for the production data, the OECD for R&D expenditure, and Eurostat and the OECD for GDP deflators.

All monetary variables in our data set 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, labor, 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 the online appendix.

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 preestimation 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 & Wu, 1999) and second generations (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 the online 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 1. The analysis of cross-section correlation indicates substantial dependence for the variables in levels as well as first differences.[16]

Aquí la casca que esto es para industria x país

V. 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 (Banerjee, Eberhardt, & Reade, 2010).[17] 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,

$$y_{it} = \alpha_i l_{it} + \beta_i k_{it} + \gamma_i r_{it} + \lambda_{it} + \psi_i + e_{it} \qquad (13)$$
$$e_{it} = \rho_i e_{i,t-1} + u_{it},$$

where l, k, and r are labor, 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 varying degrees.[18] Typical pooled estimators include the least

[16] Interestingly, testing the residuals from a pooled AR(2) regression for each of the variables cannot reject cross-section independence for value-added, labor 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.

[17] 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 (see Eberhardt, 2012); `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 web pages.

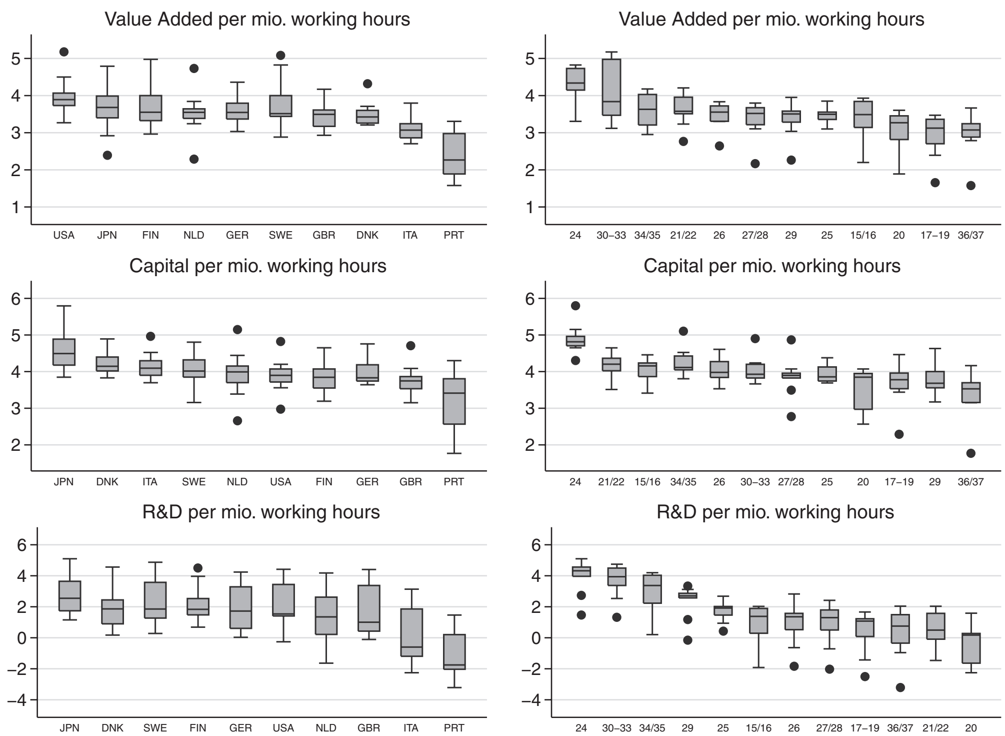
[18] Our main focus is on the most flexible setup where each country-industry can have a different set of technology parameters. Results for alternatives (country- or industry-level homogeneity) are available on request.

TABLE 3.—SUMMARY STATISTICS

	Mean	Median	SD	Minimum	Maximum
Levels					
Value-added (million euros)	27,805	7,992	52,554	290	782,206
Labor (million hours worked)	917	393	1,219	15	6,612
Physical Capital (million euro)	40,462	14,535	64340	242	459,870
R&D Capital (million 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 Labor (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 Labor	−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

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

FIGURE 1.—LABOR-DEFLATED INPUT VARIATION ACROSS COUNTRIES AND INDUSTRIES



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

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 & 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 $\rho_i \neq 0$:

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

substitute $\rho_{i,t-1}$ in 13.b and get this

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

POLS: pooled OLS (with year fixed effects); 2FE: two-way fixed effects; FD: first-difference OLS; BB: Blundell and Bond (1998); CCEP: Pooled Pesaran (2006); CCE: common correlated effects; MG: Pesaran and Smith (1995) mean group; CDMG: cross-section demeaned mean group; CMG: Pesaran (2006) CCE mean group version.

Equation (14) is commonly estimated in an unrestricted version without the nonlinear (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 micropanel estimator by Blundell and Bond (1998, BB).[19] 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 micropanel estimator has become very popular for use in macropanel data. The BB estimator solves the identification problem we discussed in section IIIB (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.[20] 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 and any other type of spillovers or common shocks. As we developed above, the various types of cross-section correlation are modeled in our empirical strategy using unobserved common factors.[21] The distinction between the left and right columns 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 and 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 that 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 while 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, Pesaran, & Yamagata, 2011).[22] To see the intuition behind this approach, consider the cross-section average of our pet model from equations (7), (8a), and (8b), replicated here for convenience,[23]

$$y_{it} = \beta_i x_{it} + \varphi_i f_t + \psi_i + \varepsilon_{it} \tag{15}$$

$$\bar{y}_t = \bar{\beta}\bar{x}_t + \bar{\varphi} f_t + \bar{\psi} \qquad \text{given } \bar{\varepsilon}_t \to 0 \text{ as } N \to \infty$$

$$\Leftrightarrow \quad f_t = \bar{\varphi}^{-1}(\bar{y}_t - \bar{\psi} - \bar{\beta}\bar{x}_t), \tag{16}$$

where cross-section averages at time t are defined as $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$ and $\bar{x}_t = N^{-1}\sum_{i=1}^{N} x_{it}$.[24] 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

$$y_{it} = \beta_i x_{it} + \varphi_i \bar{\varphi}^{-1}\left(\bar{y}_t - \bar{\psi} - \bar{\beta}\bar{x}_t\right) + \psi_i + \varepsilon_{it}, \tag{17}$$

$$\Leftrightarrow \quad y_{it} = \beta_i x_{it} + \pi_{1i}\bar{y}_t + \pi_{2i}\bar{x}_t + \pi_{3i} + \varepsilon_{it}. \tag{18}$$

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

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

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

[21] 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 such as first differencing.

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

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

[24] 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). See Eberhardt and Teal (forthcoming) for a practical application.

Not same technology or different dynamics

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 or 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.

VI. 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 and 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.

A. *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 labor 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,

Common parameter models

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

POLS: pooled OLS; 2FE: two-way fixed effects; FD: OLS with variables in first differences, CCEP: pooled Pesaran (2006) estimator. Absolute t-statistics in parentheses, constructed from White heteroskedasticity-robust standard errors. Significance at ** 5% and *** 1%.

Diagnostics: CRS: Wald test for H_0 of constant returns to scale (labor, physical capital, and R&D capital; p-values reported). AB test: Arellano and Bond (1991), 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.

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 favorable 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 report only 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 labor vary substantially across the five models presented, from .1 to .9 and −.5 to .7, respectively. All but the POLS model in column 1 result in very low 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 favorable residual diagnostics, these results are still somewhat questionable. The identification problem highlighted in equation (9) is most likely the culprit for this outcome. The poor performance of the BB estimator (negative, albeit insignificant, labor coefficient), relying on lagged levels variables as instruments for contemporaneous

These two columns
allow for cross-section dependence
(spillovers)

TABLE 6.—POOLED PRODUCTION FUNCTIONS (DYNAMIC)

	POLS (1)	2FE (2)	BB (3)	CCEP (4)	CCEP (5)
A. Long-Run Coefficients (Unrestricted Models)					
Labor	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)
B. Long-Run Coefficients (Restricted Models)					
Labor		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

BB: Blundell-Bond (1998) system GMM estimator. See table 5 for details of tests and other estimators. Absolute t-statistics in brackets, constructed from White heteroskedasticity-robust standard errors. Significance ***5% and ***1%.

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 (1)	CDMG (2)	CMG (3)	CMG (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)
Country 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

Estimators: MG: mean group, CDMG: cross-sectionally demeand MG; CMG: Pesaran (2006) common correlated effects MG. Absolute t-statistics in parentheses following Pesaran and Smith (1995). Significance at the **5% and ***1%. 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). See also table 5 for details on diagnostic tests.

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 micropanel estimator (BB) fails the instrument validity (Sargan) test.[25] Once we take the possibility of cross-section dependence (spillovers, common shocks) explicitly into account in the models in columns 4 and 5 in table 6, we see a substantial reduction in the coefficient of R&D capital, and we can no longer detect a statistically significant impact. Given their favorable diagnostics, our preferred dynamic pooled models are the standard and augmented CCEP in columns 4 and 5.

We argued in section IIIB 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 mismeasurement following Schankerman (1981). However, the data required to correct for double-counting and expensing are available only 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: directly correcting the input variables and augmenting the specification with the omitted variables. Furthermore, the CCEP estimators were dropped since their use would have led 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 the online appendix.

B. *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 labor 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

[25] 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.

TABLE 8.—HETEROGENEOUS PRODUCTION FUNCTIONS (DYNAMIC)

	MG (1)	CDMG (2)	CMG (3)	CMG (4)
Long-Run Coefficients (Restricted Models)				
Labor	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

See tables 6 and 7 for details. Absolute t-statistics in brackets, following Pesaran and Smith (1995). Significance at the **5% and ***1%. All averages reported are unweighted means. The common factor restrictions cannot be rejected in any of the four models; we therefore report only 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).

income share in developed economies (Gomme & 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 do 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 favorable 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 have an impact on 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 a more flexible structure in the CMG models.[26]

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 eight and seventeen 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 view these results only 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.[27] 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 or other unobserved shocks but econometrically raises serious concerns regarding the 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 that yield a large and statistically significant coefficient of own-R&D are seriously misspecified (nonstationary, serially correlated, or cross-sectionally dependent residuals). In contrast, once our diagnostic tests are more favorable, 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.

VII. Conclusion

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. Our main claim is that

[26] 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.

[27] 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.

the Griliches framework is inadequate even when the analysis focuses exclusively on private returns to R&D.

Using data for twelve industrial sectors in ten 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 that 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 accounts not just 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. Therefore, it cannot 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, including only measures capturing R&D spillovers in the empirical equation is unlikely to account appropriately for a 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.[28] 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.

[28] 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 recognize 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.

REFERENCES

Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti, "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association* 4 (2006), 37–74.

Arellano, Manuel, and Stephen R. Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies* 58 (1991), 277–297.

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

Azariadis, Costas, and Allan Drazen, "Threshold Externalities in Economic Development," *Quarterly Journal of Economics* 105 (1990), 501–526.

Baldwin, Richard, Henrik Braconier, and Rikard Forslid, "Multinationals, Endogenous Growth, and Technological Spillovers: Theory and Evidence," *Review of International Economics* 13 (2005), 945–963.

Banerjee, Anindya, Markus Eberhardt, and J. James Reade, "Panel Estimation for Worriers," Oxford University, Department of Economics discussion paper 514 (November 2010).

Banerjee, Abhijit V., and Andrew F. Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy* 101 (1993), 274–298.

Blundell, Richard, and Stephen R. Bond, "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics* 87 (1998), 115–143.

Bowsher, Clive, "On Testing Overidentifying Restrictions in Dynamic Panel Data Models," *Economics Letters* 77 (2002), 211–220.

Boyd, Derick, and Ron P. Smith, "Some Econometric Issues in Measuring the Monetary Transmission Mechanism, with an Application to Developing Countries" (pp. 68–99), in Lavin Mahadeva, and Peter Sinclair (eds.), *Monetary Transmission in Diverse Economies* (Cambridge: Cambridge University Press, 2002).

Cameron, Gavin, "Innovation and Economic Growth," Centre for Economic Performance discussion paper 277 (February 1996).

Cameron, Gavin, James Proudman, and Stephen Redding, "Technological Convergence, R&D, Trade and Productivity Growth," *European Economic Review* 49 (2005), 775–807.

Chudik, Alexander, M. Hashem Pesaran, and Elisa Tosetti, "Weak and Strong Cross Section Dependence and Estimation of Large Panels," *Econometrics Journal* 14 (2011), C45–C90.

Coe, David, and Helpman Elhanan "International R&D Spillovers," *European Economic Review* 39 (1995), 859–887.

Conley, Timothy, and Ethan Ligon, "Economic Distance and Cross-Country Spillovers," *Journal of Economic Growth* 7 (2002), 157–187.

Crepon, Bruno, Emmanuel Duguet, and Jacques Mairesse, "Research, Innovation and Productivity: An Econometric Analysis at the Firm Level," *Economics of Innovation and New Technology* 7 (1998), 115–158.

Cuneo, Philippe, and Jacques Mairesse, "Productivity and R&D at the Firm Level in French Manufacturing" (pp. 376–392), in Zvi Griliches (ed.), *R&D, Patents and Productivity* (Chicago: University of Chicago Press, 1984).

Doraszelski, Ulrich, and Jordi Jaumandreu, "R&D and Productivity: Estimating Production Functions When Productivity Is Endogenous," CEPR discussion paper 6636 (January 2008).

Eberhardt, Markus, "Estimating Panel Time Series Models with Heterogeneous Slopes," *Stata Journal* 12 (2012), 61–71.

Eberhardt, Markus, and Christian Helmers, "Untested Assumptions and Data Slicing: A Critical Review of Firm-Level Production Function Estimators," Oxford University, Department of Economics discussion paper 513 (November 2010).

Eberhardt, Markus, and Francis Teal, "Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics," *Journal of Economic Surveys* 25 (2011), 109–155.

——— "No Mangoes in the Tundra: Spatial Heterogeneity in Agricultural Productivity Analysis," *Oxford Bulletin of Economics and Statistics*. doi:10.1111/j.1468-0084.2012.00720.x.

Edworthy, Emma, and Gavin Wallis, "Treating Research and Development as a Capital Asset," *Economic and Labor Market Review* 1:2 (2007), 16–25.

Gomme, Paul, and Peter Rupert, "Measuring Labor's Share of Income," Federal Reserve Bank of Cleveland Policy discussion papers (November 2004).

Goto, Akira, and Kasuyuki Suzuki, "R&D Capital, Rate of Return on R&D Investment and Spillover of R&D in Japanese Manufacturing Industries," this REVIEW 71 (1989), 555–564.

Griffith, Rachel, Stephen Redding, and John van Reenen, "Mapping the two Faces of R&D: Productivity Growth in a Panel of OECD Industries," this REVIEW 86 (2004), 883–895.

Griliches, Zvi, "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics* 10 (1979), 92–116.

——— "Returns to Research and Development Expenditures in the Private Sector" (pp. 419–461), in: John W. Kendrick and Beatrice N. Vaccara (eds.), *New Developments in Productivity Measurement and Analysis* (Chicago: University of Chicago Press, 1980).

Grossman, Gene, and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *Review of Economic Studies* 58 (1991), 43–61.

Guellec, Dominique, and Bruno van Pottelsberghe, "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* 66 (2004), 353–378.

Hall, Bronwyn, Jacques Mairesse, and Pierre Mohnen, "Measuring the Returns to R&D," NBER working paper 15622 (December 2009).

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

Hall, Bronwyn, and Jacques Mairesse, "Exploring the Relationship between R&D and Productivity in French Manufacturing Firms," *Journal of Econometrics* 65 (1995), 263–293.

Holly, Sean, M. Hashem Pesaran, and Takashi Yamagata, "A Spatio-Temporal Model of House Prices in the USA," *Journal of Econometrics* 158 (2010), 160–173.

Kao, Chihwa, "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data," *Journal of Econometrics* 65 (1999), 9–15.

Kapetanios, George, M. Hashem Pesaran, and Takashi Yamagata, "Panels with Nonstationary Multifactor Error Structures," *Journal of Econometrics* 160 (2011), 326–348.

Keller, Wolfgang, "Are International R&D Spillovers Trade Related? Analyzing Spillovers among Randomly Matched Trade Partners," *European Economic Review* 42 (1998), 1469–1481.

——— "Trade and the Transmission of Technology," *Journal of Economic Growth* 7 (2002a), 5–24.

——— "Geographic Localization of International Technology Diffusion," *American Economic Review* 92 (2002b), 120–142.

——— "International Technology Diffusion," *Journal of Economic Literature* 42 (2004), 752–782.

Maddala, G. S., and Shaowen Wu, "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test," *Oxford Bulletin of Economics and Statistics* 61 (1999), 631–652.

Nickell, Stephen, "Biases in Dynamic Models with Fixed Effects," *Econometrica* 49 (1981), 1417–1426.

Pesaran, M. Hashem, "General Diagnostic Tests for Cross Section Dependence in Panels," Cambridge working papers in economics 0435 (June 2004).

Pesaran, M. Hashem, "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica* 74 (2006), 967–1012.

——— "A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence," *Journal of Applied Econometrics* 22 (2007), 265–312.

Pesaran, M. Hashem, and Ron P. Smith, "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics* 68 (1995), pp. 79–113.

Romer, Paul, "Endogenous Technological Change," *Journal of Political Economy* 98 (1990), 71–101.

Roodman, David, "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics* 71 (2009), 135–158.

Schankerman, Mark, "The Effects of Double-Counting and Expensing on the Measured Returns to R&D," this REVIEW 63 (1981), 454–458.

van Pottelsberghe, Bruno, and Frank Lichtenberg, "Does Foreign Direct Investment Transfer Technology across Borders?" this REVIEW 83 (2001), 490–497.