Research Policy 44 (2015) 381–393

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

The production function of top R&D investors: Accounting for size and sector heterogeneity with quantile estimations

Sandro Montresor[b], Antonio Vezzani[a,*]

[a] JRC-IPTS, European Commission, Seville, Spain
[b] Department of Economics, University of Bologna, Italy

ARTICLE INFO

Article history:
Received 17 July 2013
Received in revised form 15 May 2014
Accepted 18 August 2014
Available online 4 September 2014

JEL classification:
D24
D21
O30

Keywords:
Production function
R&D
Firm and sector heterogeneity

ABSTRACT

This paper aims at showing how quantile estimations can make the analysis of the firm's production function better able to deal with the innovation implications of production. In order to do this, we provide evidence of how top world R&D investors differ in the production impact of their inputs and in their rate of technical change. We use the EU Industrial R&D Investment Scoreboard and carry out a quantile estimation of an augmented Cobb–Douglas production function for a panel of more than 1000 companies, covering the 2002–2010 period. The results of the pooled sample are contrasted with those obtained from the estimates for different groups of economic sectors. Returns to scale are bounded by the size of the firm, but to an extent that decreases with the technological intensity of the sector. The output return of knowledge capital is the largest, irrespective of firm size, but in high-tech sectors only. Elsewhere, physical capital is the pivotal factor, although with size variations. The investigated firms also appear different in their technical progress: embodied in mid-high and low/mid-low tech sectors, and disembodied in high-tech sectors.

1. Introduction

Confirming what Griliches and Mairesse noticed nearly twenty years ago, the production function nowadays appears as "a tool, a framework for answering other questions, only partially related to [it]" (Griliches and Mairesse, 1995, p. 2). In innovation studies, this is proved by the extensive use of the so-called "knowledge production function" (Pakes and Griliches, 1984). Its relationship with its "standard" counterpart is at most partial and its "eclectic" nature apparently makes its application less problematic. On the one hand, the emphasis on a theoretically consistent "transformation" logic (of inventive inputs into innovative output) ends up taking some focus away from the burden of stringent hypotheses (e.g. in terms of returns to scale), which often accompany its application to the production realm. On the other hand, the primer role of feed-backs in the innovation process (e.g. from profitability to innovation investment) makes the search for an exogeneous treatment of the relevant (knowledge) inputs (Crepon et al., 1998) more fundamental than it is in the production analysis, while rendering the actual functional specification of their output (innovation) impact relatively less crucial.

Benefiting from an apparently higher tractability, estimates of the knowledge production function have been growing in number in the last thirty years and have substantially contributed to opening up the innovation "black-box". Conversely, estimates of the production function as such for the same purpose have been surprisingly under-used. Given the relevance of production-related issues in innovation, and their policy and strategic impact, this seems quite unfortunate. For example, the marginal productivity of different production inputs (labour and capital, to start with) is an important element in detecting a firm's opportunities for "embodied" vs. "disembodied" technological change. Similarly, the evidence of increasing, constant or decreasing returns to scale can inform us about the firm's capacity to overcome the indivisibilities that often hamper the achievement of significant innovation outcomes.

Although quite important, analysis of these innovation implications has been hampered by the burden of well-known conceptual hypotheses and econometric problems that affect estimation of the production function. In addition to these, we retain that a crucial obstacle to the simple use of the production function is represented by the inner heterogeneity that firms have been found to show in innovation, in particular with respect to their size and specific economic sector. This is a basic pillar of modern economics

* Corresponding author.
E-mail addresses: sandro.montresor@unibo.it (S. Montresor), antonio.vezzani@europa.ec.eu (A. Vezzani).

http://dx.doi.org/10.1016/j.respol.2014.08.005

of innovation: firms of different sizes show inherently diverse capacities of introducing, appropriating and exploiting innovations, and additional heterogeneity comes from the industrial structure and dynamics of their different sectors of activity (Cohen, 2010). Accordingly, this heterogeneity should be also retained in investigating firms' capacity to turn their production inputs into (production) output.

While there are ways to integrate them in order to account for this heterogeneity, econometric estimates of a parametric kind are not fully equipped to accurately illustrate its impact on production and its relevance for innovation. However, less standard semi-parametric techniques can be used to this scope and interesting implications can be obtained from them. This is particularly the case for *quantile regression*, on which the present paper focuses, aiming to show that it can be a useful analytical tool for a micro-econometric estimate that directly tackles firms' heterogeneity and thus increases its utility in the analysis of innovation. In particular, the quantile estimation can help us detect how much the production impact of firms' inputs varies along different firm-size quantiles and in different economic sectors.

We carry out this estimate on a sample of more than 1000 top R&D investors – representing nearly 80% of total world R&D – over the 2002–2010 period. Their high R&D intensity makes them a sample of firms with substantial innovative efforts (highly innovative, if we use an input kind of proxy for innovation) and with a relatively homogenous pattern of innovation (i.e. relying on internal and formal innovative efforts). Furthermore, the ranking criterion with which the sample is built up leads it to be dominated by large-sized (or, at the least, medium-sized) companies. Given these common features, one could argue that their production behaviour and performance are relatively homogeneous and that any policy supporting them may well require a similar kind of approach. These considerations make our search for heterogeneity in the production function of these firms – both in terms of size and sector of economic activity – particularly interesting. Should we actually find traces of this heterogeneity, the exercise that we propose would become even more compelling for a more general kind of sample.

The remainder of this paper is organised as follows. Section 2 outlines the literature from which our analysis draws. Section 3 illustrates the data and econometric methodology. Section 4 reports and discusses the results. Section 5 concludes with a set of policy implications.

2. Theoretical background

The so-called "knowledge production function", introduced for the first time by Pakes and Griliches (1984), is certainly the most used and debated form of a relation between inputs and outputs in innovation studies. Inspired by Griliches' (1979, 1998) incessant search for a nexus between R&D and productivity, such a production function is based on the informative value of patents, as a proxy of the firm's capacity of transforming R&D into economically valuable knowledge. Its applicability to the analysis of the firm's innovation, in particular under the limitations of cross-section data, has been substantially increased by the now popular Crepon et al. (1998) (CDM) model, in which the (R&D) inputs of the same function are recursively endogenized on the basis of theoretical and empirical insights. More recently, the same function has been further refined by disentangling the different nature of its knowledge inputs (e.g. Zucker et al., 2007; Ramani et al., 2008; Czarnitzki et al., 2009) and outputs (e.g. Evangelista and Vezzani, 2010, 2012). Its extension to the inclusion of R&D/knowledge spillovers, especially with respect to co-localised (e.g. regional) firms, represents an additional path along which the 'augmented' knowledge production function has recently been developed (e.g. Ponds et al., 2010; D'Agostino et al., 2013).

Less popular in innovation studies, however, is the analysis of the firm's production function as such, that is the technical relationship between the firm's production inputs (e.g. labour and capital) and output (i.e. goods and services). As is well-known, this neglect is mainly due to the famous critique of a "black-boxed" (neoclassical) view of technological change (Rosenberg, 1982, 1994), of which the production function represents the main conceptual pillar.[1] Additional resistance to its use, in innovation studies as well as elsewhere, can be found in problems pertaining to an "econometrically determined production function" (Shephard, 1974, p. 407) and the delicate questions posed by its estimation (Griliches and Mairesse, 1995), starting with its identification and with the endogeneity (simultaneity) of its inputs (input–output relationship).[2]

These theoretical and empirical issues surely downplay the utility of the production function in dealing with innovation. However, their incidence can be attenuated, and utility thus increased, when a different approach is adopted in determining its estimation, as we suggest in the following section. Indeed, the search for a more "innovation-friendly" use of the production function would seem recommendable, given the important aspects of innovation to which it directly and indirectly refers. First of all, the production function has offered important insights in analysis of the so-called "embodied technological change", as distinct from disembodied technological change. Following Solow's (1960) seminal idea, according to which TFP changes can be due to new (better or higher quality) vintages of capital goods, physical investments have attracted a lot of attention as an innovation driver, even with respect to the existing firm's techniques (Salter, 1960).[3] Investment-specific technological change was initially linked to the ICT revolution in the early 1970s, and to the incorporation of computerised machinery in production (e.g. Yorukoglu, 1998). More recently, however, its role has been more generally argued in other sectors as well, where (R&D) knowledge capital is apparently the main driver (e.g. Ortega-Argilés et al., 2011). This has led to important policy implications about how to spur innovation and productivity, especially in those portions of the European area where industrial structure is still rooted into capital-intensive specialisation models (Conte and Vivarelli, 2005).

On this basis, in innovation studies it becomes extremely important "going-back" to the production impact of physical capital, vs. labour and knowledge capital, in the way in which the estimate of the production function allows us to do this. The search for more accurate way of carrying out the estimate is thus compelling.

A similar need emerges from the centrality that through the use of the production function, scholars have long found for returns

[1] As is also well-known, other limitations in its general use are related to the problematic way capital is encapsulated within it – the famous "Cambridge capital controversy" (see Cohen and Harcourt, 2003) – and to the absence of an actual entrepreneurial decisional process that the function entails – as in the Austrian theory (see Foss, 1994).

[2] These problems have attracted most econometric attention for the production function in the last fifteen years, along which the use of instrumental variables and of fixed effect estimation has been enriched by the dynamic panel literature (following Arellano and Bond (1991)) and by that on observed input decisions (following Olley and Pakes (1996)).

[3] Originally explained with quite strong hypotheses – e.g. by assuming a constant rate of technical change and treating consumption and investment goods as perfect substitutes in production (Solow, 1960) – embodied technical change was subsequently accounted for in more general terms (Jorgenson, 1966) and more recently inserted into a general equilibrium framework, through vintage-capital models in which production becomes more efficient over time (e.g. Greenwood et al., 1997). In empirical terms, its estimates were mainly carried out with a "price-approach", claiming that official price indices do not accurately capture changes in the quality of equipment, and that their comparison with more articulated price indexes (e.g. that by Gordon (1990)) can be taken as a proxy of embodied technical changes.

to scale in accounting for productivity growth (Brown and Popkin, 1962). As is well known, the “switch” from constant to increasing returns to scale has been one of the conceptual leverages through which growth has been endogenised in “new growth theories”: not only by considering the R&D sector (Aghion and Howitt, 1992), but also the special nature of some firms’ production inputs (Romer, 1990). Accordingly, a broader and possibly more accurate refocusing on returns to scale in production appears desirable: above all in guiding European policy makers in support of their exploitation in mid- and mid-low technological sectors/countries (Heidenreich, 2009; Santamaría et al., 2009).

Finally, although this could appear at odds with the most recent efforts to open the “black-box” of technological change (Dosi, 1988), the estimate of the production function can be useful to get a first quantitative “taste” of the rate at which the knowledge available to the firm grows over time. If properly retained, such an indication could be used as a reference for other, less black-boxed measurements of the firm’s knowledge production function.

All of the previous arguments call for a refocusing on the “old” production function. However, as we said, this should retain the essential elements of heterogeneity that characterise innovation, and that could also be expected in the production realm.

As far as R&D and innovation are concerned, the heterogeneity shown by firms of different sizes and sectors dates back at least to the work of Joseph Schumpeter in the previous century. The subsequent debate on ‘technological regimes’ of Schumpeter Mark I and Mark II kind has provided us with new insights on this issue (e.g. Breschi et al., 2000). Along the same research-line, important results have been obtained by the literature on “sectoral innovation systems” (Malerba, 2002). Those firms whose inventive output belongs to different technological classes (e.g. in terms of patents) have been shown to follow different innovative patterns – for example, “widening” (à la Schumpeter Mark I), rather than “deepening” (à la Schumpeter Mark II) – above all in terms of firm-size, market concentration and industry dynamics (Malerba and Orsenigo, 1995, 1996). Quite interestingly, these size and sectoral elements of heterogeneity in innovation hold true across different countries, whose specificities “just” introduce differences in innovative patterns, within each and every technological class. These sectoral patterns of technological activity – and of R&D intensity, in particular – appear important in driving the dynamics of countries’ market shares on the global level. Their evolutionary dynamics should thus be the starting point for building up proper models of structural change and aggregated productivity growth (Montobbio, 2002, 2003).[4]

Although it has received less attention, substantial heterogeneity should also be expected by looking at the production function that innovative firms of different sizes and sectors use in employing their inputs for obtaining their production (rather than innovative) output. First of all, firms of different sizes could benefit (suffer) from returns to (diseconomies of) scale to a different extent. The standard (i.e. labour-capital based) micro-economic argument would suggest that smaller firms are better placed to benefit from increasing returns to scale, whereas larger ones could suffer from decreasing returns due to technical inefficiencies and/or managerial costs. However, in firms which heavily invest in innovation – like those top R&D investors that we investigate – the crucial role that knowledge capital plays, especially in relation to an increase in their scale of operation, could alter this picture. This is a quite well-established argument in industrial studies (Scherer, 1965; Acs and Audretsch, 1987), which the results of the new growth theories about R&D spillovers and returns to scale (e.g. Aghion and Howitt, 1992) have reinvigorated. Furthermore, the techno-economic features of the sectors in which the firms operate – and their intensity of physical and knowledge capital, in particular – could introduce differences in the way returns to scale emerge along their size distribution. Finally, a differentiating impact on the characterisation of returns to scale can be exerted by the different stages of their technology/product development (e.g. Utterback and Abernathy, 1975).

A second point concerns the marginal returns of the factors that firms use in production, which are also supposedly size- and sector-specific. For example, the indivisibilities to which capital investments are generally exposed (Tone and Sahoo, 2003) would suggest that, compared to that of labour, their production impact is higher in larger rather than in smaller firms. However, in firms that largely invest in innovation, the marginal contribution of knowledge capital is also expected to play an important role and show a different impact at different size levels (Lööf and Heshmati, 2002). On the one hand, by spreading the outcome of their projects over a larger level of output, bigger firms could be expected to have higher returns from R&D (Cohen and Klepper, 1996). On the other hand, smaller firms may benefit from more creative R&D projects and have a more technical scope for their exploitation (Acs and Audretsch, 1987). Once again, sector specificities matter here as well. For example, following Cohen and Klepper (1996), the relationship between R&D and size should be weaker in industries where innovation may lead to stronger growth or where innovations are of a more disembodied form.

Last but not least, in spite of the constraints that the estimate of the production function impose on this kind of detectable technical change (which we will discuss in the next section), its rate is expected to be variable along the observed distribution of firms and to show differences across sectors as well. Although only indirectly, this is suggested by the emerging studies on the heterogeneity of the innovative output of manufacturing firms and of their patterns of economic growth (Ciriaci et al., 2012; Coad and Rao, 2006, 2008).

All in all, the heterogeneity that firms show in production-related issues appears conceptually significant. Unfortunately, available evidence, mainly based on parametric techniques, provides us with only scattered support to its relevance. In order to offer more general insights, in the next section we propose and carry out an empirical application which presents firms’ heterogeneity in production more systematically by using the quantile regression approach. In so doing, we extend previous applications of the quantile approach to the “frontier production framework”, whose aim is, differently from ours, to address heterogeneity in production efficiency in general with respect to a standard set of inputs (that is, in absence of knowledge capital).[5] Furthermore, we also contribute to extending the focus of existent literature on quantile regression analyses which, in line with the “conceptual-bias” we argued in Section 1, has also recently concentrated on the knowledge “counterpart” of the production function, by addressing CDM-like relationships between R&D, innovation and economic performance (e.g. Nahm, 2001; Coad and Rao, 2006, 2008; Kaiser, 2009; Stam and Wennberg, 2009; Ebersberger et al., 2010; Segarra and Teruel, 2011; Falk, 2012; Mata and Wörter, 2013).[6]

[4] Important elements of analysis have also emerged from the specific literature on the role of market structure for R&D and innovation (e.g. Kamien and Schwartz, 1982). Size and sector specificities have also been identified by looking at innovation diffusion among firms. From the seminal Pavitt taxonomy (Pavitt, 1984), up to the most recent sectoral classification in terms of innovation (Castellacci, 2008), the differences that firms show in terms of internal and external knowledge sources, technology transfer, and innovation strategies (to mention a few) have also (although not uniquely) been related to their size and to the techno-economic characteristics of their sector of activity.

[5] See, for example, Bernini et al. (2004).

[6] Similarly, we also contribute to generalising previous quantile regression studies on the returns of skills for different levels of wages/earnings and on the returns of

3. Empirical application

3.1. Data

We estimate the production function of a sample of firms contained in the EU Industrial R&D Investment (IRI) Scoreboard (http://iri.jrc.ec.europa.eu/). This is a scoreboard analysis of top R&D investors, in Europe and in the rest of the world, which the Institute of Prospective Technological Studies (IPTS, Joint Research Centre, European Commission) has conducted annually since 2004. Data for the relative annual ranking is collected from the companies' latest published accounts (e.g. 2012 for the Scoreboard of 2013), by referring to the ultimate parent company in the case of consolidated groups. Companies which are subsidiaries of any other company are not listed separately, while subsidiaries are included when consolidated group accounts of the ultimate parent company are not available. The focal variable of the Scoreboard is the companies' cash investment in R&D (as from international accounting standards) funded by the companies themselves, excluding those undertaken under contract for customers such as governments or other companies. In addition, data on net sales, operating profit, capital expenditure and number of employees is reported, though with variable coverage (for more details, see IRI-IPTS, European Commission, 2013). The IRI Scoreboard comprehends nearly all the big players of the R&D investments in the World (especially in mid-high and high-tech sectors) and accounts for nearly 80% of this total expenditure. The sample selection bias in investing the relative population can thus be deemed not that large. Due to these features, the sample has been previously used for investigating a number of interesting issues pertaining to the same population (e.g. Cincera and Ravet, 2010; García-Manjóna and Romero-Merino, 2012).

By integrating the yearly Scoreboards with other data from IRI sources (in particular, integrations/checks carried out by its researchers through examining the companies' web-sites), and by merging them, we have obtained a panel of 1024 companies, over the 2002–2010 period.[7]

The sample is made up of large companies (28,016 employees on average), which however show appreciable size variations across different sector groups. Firms in high-tech sectors (i.e. with an R&D intensity higher than 5%)[8] are comparatively smaller (14,835 employees on average) than those in medium/high-tech sectors (R&D intensity between 2% and 5%, with 32,048 employees on average) and medium/low ones (R&D intensity lower than 2%, with 48,386 employees on average) (Tables A1 and A2). Size heterogeneity is also relevant within sectors. The within-sector standard deviation of employment is appreciable (38,942, 54,910, and 77,820, for the three sector groups) and median values are much lower than their respective mean averages (3034, 11,821, and 21,742, respectively). The groups of sectors that we have identified in terms of R&D intensity are also heterogeneous when we look at the different economic activities that they encompass (Table A2). However, although with some degree of approximation (mainly due to the firms' sizes), the technological base that they share can be traced back to that of the Pavitt (1984) taxonomy. All of these elements will have to be considered in interpreting the results of our empirical application.

3.2. Variables and econometric strategy

Following the bulk of the literature, for the sake of analytical tractability and ease of interpretation, we adopt a Cobb–Douglas formulation for the production function of firm i at time t, augmented to include R&D-based "knowledge capital", that is:

$$Y_{it} = A_t K_{it}^{\alpha} RD_{it}^{\beta} L_{it}^{\gamma} e^{u_{it}} \tag{1}$$

Y denotes the firms' production output, L stands for labour, K and RD for physical and knowledge capital stocks, respectively. A_t represents the technology in use and is defined as $A_t = Ae^{\rho t}$, where t is the time index and u_{it} represents the systematic component of the unmeasured factors, assumed to be randomly distributed. α, β, γ, and ρ are the parameters of interest.

As is well-known, the Cobb–Douglas production function is the only linearly homogeneous function that entails constant factor shares (or marginal rates of return) and a unitary elasticity of substitution: two hypotheses that are hardly satisfied in empirical applications. Although an intrinsic limitation, we have opted to stick to it – a price to pay in order to illustrate, in an intuitive way, the kind of heterogeneity (i.e. in terms of size and sector) we are interested in.[9]

In Eq. (1), Y is measured in terms of firm's turnover.[10] L is obtained by considering the number of its employees, K and RD are built up using the perpetual inventory method (Hall and Mairesse, 1995). For each firm i operating in a certain sector m, at time t the relevant *Stock* is defined by the following formulas:

$$Stock_{jit=2002} = \frac{I_{ji=2002}}{\bar{g}_m + \delta_j} \quad \text{for} \quad t = 2002 \tag{2}$$

$$Stock_{jit} = Stock_{jit-1}(1-\delta_j) + I_{jit} \quad \text{for} \quad t > 2002$$

where t = 2002,. . .,2010. For each kind of *Stock* (j = K; RD), I represents the relative investments observed in the sample, $\bar{g}$ is their sectoral average growth rate, and δ is the depreciation rate of capital. Following the extant literature (Hall and Mairesse, 1995), δ has been set to 15% for knowledge and 8% for physical capital, respectively (see Section 4.1 for a robustness check of these values).

Taking the logarithms of (1), we get the following estimation equation, where small letters stand for logarithms:

$$y_{it} = a + \rho t + \alpha k_{it} + \beta rd_{it} + \gamma l_{it} + u_{it} \tag{3}$$

human capital for different productivity levels (for a review, see Bartelsman et al., 2013).

[7] The panel is slightly unbalanced, due to the fact that some of the current top R&D investors were not present in the ranking in earlier years (e.g. HTC). As for the demographic industrial evolution of the panel, in the case of a demerger, the full history of the continuing entity is included. In order to avoid double counting problems, the history of the demerged company can only go back as far as the date of the demerger. In case of M&As, pro-forma figures for the year of acquisition are used along with pro-forma comparative figures if available. For extra-Euro companies, currency amounts have been translated at the Euro exchange rates ruling at the latest day of the previous year, and the exchange rate conversion has also been applied to the historical data. While in so doing the Scoreboard reflects the domestic currency results of the companies, rather than economic estimates of current purchasing parity results, this fact does not have an impact on the kind of estimates on which we focus. Finally, all the relevant figures have been deflated using the GDP deflators published by the World Bank, and using 2002 as the reference year. For companies located in the Cayman Islands we applied the World average deflator. In the case of companies based in Taiwan (Chinese Taipei), we used the "Implicit GDP Price Indices" taken from the OECD-MSTI database instead.

[8] Consistently with the IRI Scoreboard, R&D intensity is here defined as the ratio between R&D investments and turnover. The threshold values for identifying sector groups are also drawn from the IRI Scoreboard.

[9] A more flexible functional form, among those which are used in micro-econometric estimations (Battese and Broca, 1997), while remedying the flaws of the Cobb-Douglas production function, does not have the same advantages of analytical tractability we are able to exploit with this function (Douglas, 1976).

[10] The choice of using the firm's gross revenues (i.e. gross output) rather than its value-added as a proxy of its output has been mainly driven by issues of data availability, in particular, by the lack of systematic figures on the cost of labour for non-EU based companies. While value-added based estimates would have possibly led to different results, both the methodologies have been shown to have pros and cons (e.g. Cobbold, 2003; Mairesse and Jaumandreu, 2005).

A list of dummy variables, at the industry (ICB, Industry Classification Benchmark, 4-digit level), time and country levels, is included in the estimation.

Consistently with the use of the Cobb–Douglas functional form, the parameter ρ of Eq. (3), which captures output variations over time that are not accounted for by changes in the use of inputs, is taken to measure the firm's rate of technical progress. More precisely, ρ captures that "portion" of the firm's technological change, which is not (endogenously and directly) accounted for by the disembodied knowledge the firm creates with its R&D investments and by the embodied knowledge it encapsulates into new vintages of physical capital. The inclusion of industry, country and, above all, time controls, enables us to be confident that such a linear trend actually captures the (constant) technological shift experienced by the focal firm over time, because of the joint efforts of all the players of its industry or, more broadly, by the improvement of the technological opportunities of its sectoral system of innovation.

Eq. (3) is estimated with a quantile model – discussed below – and the relative results are compared with those obtained using three other standard approaches: (1) Ordinary Least Squares (OLS), (2) Panel Random Effects (RE), and (3) System GMM (SYS-GMM).

Among the possible alternatives, as usual, OLS is taken to represent a sort of benchmark estimate. RE, on the other hand, has been chosen in order to have a specification comparable to that of the focal quantile in terms of controls, given that the Hausman test did not provide evidence for supporting an alternative fixed effect model. Finally, we present the results from the generalised method of moments (GMM) estimators, which has become a standard methodology in the dynamic panel framework. In particular, we use the application of the GMM proposed by Blundell and Bond (1998, 2000), which is essentially built as a system (SYS) of equations, where lagged first-differences are used as instruments for equation in levels, and lagged levels as instruments for equations in first-differences.[11] The SYS-GMM framework allows us to relax the assumption of strict exogeneity of the regressors and to consider the production inputs as predetermined.

In comparison with these alternative models, the quantile model has some important properties with respect to the issue at hand (Koenker and Bassett, 1978; Koenker and Hallock, 2001). First of all, it is robust against outliers and non-normal distributed errors. Second, it allows us to estimate different measures of central tendency and statistical dispersion. Furthermore, and of greater relevance for our topic, it gives a more comprehensive picture of the relationship between variables, by directly accounting for firms' heterogeneity across the sample. Indeed, the way heterogeneity is accounted for by the quantile approach is substantially different from the other models. As is well-known, OLS estimations simply assume that unobserved heterogeneity exclusively derives from sector-, time-, and country-specific factors. The RE approach, conversely, assumes that there is an important source of heterogeneity coming from time-invariant, firm-specific factors, which can be accounted for by the idiosyncratic part of the error term (i.e. in Eq. (3), instead of estimating $a + u_{it}$, we estimate $a_i + u_{it}$).[12] Unlike the aforementioned models, the quantile approach directly controls for that part of the firms' heterogeneity that derives from sector and country-specific factors and explicitly models it in terms of independent variable levels. In brief, the parameters in Eq. (3) are allowed to vary across the firm distribution in terms of the dependent variable, that is the firms' turnover. Accordingly, an important part of the firms' heterogeneity within a specific sector (and country) is taken to derive from their size in terms of turnover.[13]

In analytical terms, we are interested in estimating $Q_\tau(y_{it}|x_{it}) = x_i'\beta$, that is the τ^{th} conditional quantile of y_{it} given x_{it}. This can be done by solving the following problem[14]:

$$\hat{\beta}_\tau = arg\,\min \sum_{i=1}^{n} \rho_\tau(y_{it} - x_{it}'\beta) \quad \text{where } \rho_\tau = u_{it}(\tau - \mathbf{1}_{(u_{it}<0)})$$

By increasing τ continuously, from 0 to 1, it is possible to trace the entire distribution of y, conditional on x (our RHS variables).

4. Results

The results of the quantile estimation provide us with interesting insights about some important issues raised by the production function analysis.

The first issue is the analysis of returns to scale, measured by the extent to which a firm's production output varies with respect to the same joint variation of all its inputs. As is well-known, depending on the former being more, equally, or less than proportional to the latter, these returns are said to be increasing, constant or decreasing, respectively. Benefiting from the properties of the Cobb–Douglas production function, we tested for whether the sum of the coefficients attached to the production factors is statistically different from 1 and looked at its actual value.[15]

Compared to standard estimates, which suggest that returns to scale are generally constant (in the case of SYS-GMM, the test relies on the relatively higher standard errors attached to the estimated coefficients), the quantile estimate points to important elements of heterogeneity in their specification (Table 1).[16]

First of all, when we consider the entire distribution of the observed firms in terms of turnover, and we pool together firms of different sectors, evidence of decreasing returns is found at the top of the distribution. Although average-based estimators hide this result, some "few" quantiles of the investigated top R&D spenders (the largest 25% of them) appear to have overcome their minimum efficient scale of production. Consistently with standard microeconomic arguments, this result holds true for the largest firms of the whole turnover distribution, while for initial and intermediate quantiles we find evidence of increasing returns to scale. Interestingly, the distribution of the whole sample 'mimics', although with a right-hand side skewness and in absence of constant ones, the patterns that returns to scale display in textbooks with respect to the production quantities of the representative firm.

This first result reveals important specifications when we look at returns to scale for different quantiles of firms within different groups of sectors (Table 2).

[11] This type of estimator is particularly suitable when, as in our case, there is a large number of panels (companies) and a few temporal observations, the explanatory variables are not strictly exogenous, and fixed effects and autocorrelation are both present within panels. In the production function framework, this approach has proved to yield more reliable parameter estimates than classical GMM estimators (Blundell and Bond, 1998, 2000).

[12] For the sake of completeness, a fixed effect approach would consider the heterogeneity as completely determined by firm-specific factors, not allowing for the inclusion of additional time-invariant controls (e.g. sectors, time and country dummies).

[13] For the sake of brevity, when we talk of firms of different size, we will implicitly refer to firms of different quantiles of their turnover distribution.

[14] $\mathbf{1}$ denotes the indicator function. For the sake of illustration, the case of $\tau = ½$ corresponds to the median regression which minimises the sum of absolute residuals, while for $\tau = 0.25$ the weighted sum of residual is minimised with weights equal to τ when residuals are negative and $(\tau - 1)$ when residuals are ≥ 0.

[15] Constant returns to scale hold when the null hypothesis is not rejected, whereas increasing and decreasing returns hold when the null hypothesis is rejected and the sum of the coefficients is greater and smaller than 1, respectively.

[16] In looking at and interpreting the estimated coefficients, it should be noted that they give information about the marginal changes that do not move an observation from its current quantile to another quantile of the distribution.

Table 1
Production function estimates – all sample.

	OLS	RE	SYS-GMM	Quantile				
				10%	25%	Median	75%	90%
Knowledge capital	0.175***	0.101***	0.176***	0.213***	0.216***	0.188***	0.164***	0.150***
	(0.006)	(0.011)	(0.058)	(0.006)	(0.008)	(0.007)	(0.013)	(0.014)
Physical capital	0.187***	0.243***	0.227**	0.160***	0.153***	0.188***	0.216***	0.219***
	(0.007)	(0.012)	(0.109)	(0.014)	(0.009)	(0.009)	(0.012)	(0.015)
Employment	0.642***	0.678***	0.436***	0.656***	0.648***	0.629***	0.602***	0.591***
	(0.008)	(0.010)	(0.029)	(0.012)	(0.012)	(0.011)	(0.011)	(0.014)
Time trend	0.015***	0.0163***	0.015***	0.006*	0.007**	0.011***	0.017***	0.018***
	(0.003)	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)
Constant	4.859***	4.477***	5.216***	3.220***	4.426***	5.105***	5.504***	5.935***
	(0.073)	(0.183)	(0.462)	(0.356)	(0.113)	(0.120)	(0.092)	(0.119)
Returns to scale	*Constant*	*Constant*	*Constant*	*Increasing*	*Increasing*	*Increasing*	*Decreasing*	*Decreasing*
Sectorial dummies	*Significant*	*Significant*	–	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*
Country dummies	*Significant*	*Significant*	–	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*
Time dummies	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*
Observations	8990	8990	7873	8990	8990	8990	8990	8990
R-squared[a,b]	0.940	0.939	*0.000*	0.786	0.792	0.791	0.779	0.756
m1 (m2)[c]			0.00 (0.27)					
Sargan[d]			0.069					

Bootstrapped standard errors in parentheses (100 replications). ICB Industrial dummies (computed at a 4-digit level) and country dummies have been tested for their joint significance at a minimum 5% level. Returns to scale have been tested from regressions estimates. To improve the readability of the table, for the SYS-GMM only the minimum distance parameters are reported.

[a] Pseudo R-square is reported for quantile estimates.
[b] For the SYS-GMM the p-value of the Wald Chi-square test is reported.
[c] Arellano–Bond test for serial correlation in the first-differenced errors.
[d] Sargan test of overidentifying restrictions.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

On the one hand, in high-tech sectors, the case of decreasing returns disappears even from the largest portion of the relative size distribution. Such a distribution reveals at worst constant returns, after showing increasing returns up to the median. A similar pattern holds true for firms operating in the mid-high tech sectors, in which, however, firms switch from increasing to constant returns at a lower tail of the relative size/turnover distribution. On the other hand, in the low/mid-low tech sectors, we do not detect increasing returns at all, not even for the smallest firms. Conversely, the largest firms in these sectors appear to be the ones which account for the evidence of decreasing returns to scale that we have found above.

If we combine this last piece of evidence with the descriptive statistics of the sample (Table A1), an interesting general result emerges. Sector-specific levels of technology and firm size intertwine in determining technical constraints to firm growth. Moving from low- to high-tech sectors, technological knowledge makes constraints on returns to scale less stringent, while the smaller firms of the sample become more equipped to benefit from them. In brief, top R&D spenders are not associated to the same opportunities of economies of scale in producing their goods/services: in particular, smaller spenders are connected with these opportunities (differently from the larger ones) when they operate in high-tech rather than low-tech sectors.

A second set of results of our estimates concerns the marginal returns of the different inputs that firms use in production. The analysis of their output elasticity provides us with some important insights. First of all, in this case as well, standard (average-based)

Table 2
Production function estimates by technological sector – quantile regression.

	High tech (HT)			Medium-high tech (MHT)			Low & medium-low tech (LMLT)		
	25%	50%	75%	25%	50%	75%	25%	50%	75%
Knowledge capital	0.296***	0.291***	0.268***	0.168***	0.146***	0.115***	0.180***	0.166***	0.110***
	(0.010)	(0.016)	(0.017)	(0.007)	(0.009)	(0.012)	(0.018)	(0.021)	(0.036)
Physical capital	0.026**	0.059***	0.103***	0.205***	0.242***	0.266***	0.370***	0.375***	0.334***
	(0.012)	(0.020)	(0.017)	(0.017)	(0.012)	(0.020)	(0.023)	(0.025)	(0.021)
Employment	0.719***	0.673***	0.621***	0.647***	0.618***	0.615***	0.449***	0.421***	0.465***
	(0.015)	(0.011)	(0.017)	(0.015)	(0.013)	(0.020)	(0.030)	(0.033)	(0.035)
Time trend	0.004**	0.007**	0.011**	0.007**	0.012***	0.018***	0.009	0.009*	0.015**
	(0.002)	(0.003)	(0.004)	(0.003)	(0.004)	(0.005)	(0.006)	(0.005)	(0.007)
Constant	3.415***	3.680***	4.238***	3.357***	3.470***	3.928***	3.429***	4.323***	5.674***
	(0.146)	(0.122)	(0.181)	(0.091)	(0.143)	(0.096)	(0.255)	(0.303)	(0.222)
Returns to scale	*Increasing*	*Increasing*	*Constant*	*Increasing*	*Constant*	*Constant*	*Constant*	*Decreasing*	*Decreasing*
Sectorial Dummies	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*
Country Dummies	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*
Time Dummies	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*	*Significant*
Observations	3621	3621	3621	3773	3773	3773	1596	1596	1596
Pseudo R-squared	0.743	0.762	0.764	0.797	0.791	0.780	0.743	0.729	0.699

Bootstrapped standard errors in parentheses (50 replications). ICB Industrial dummies (computed at a 4-digit level), time and country dummies have been tested for their joint significance at a minimum 5% level. Returns to scale have been tested from regressions estimates.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

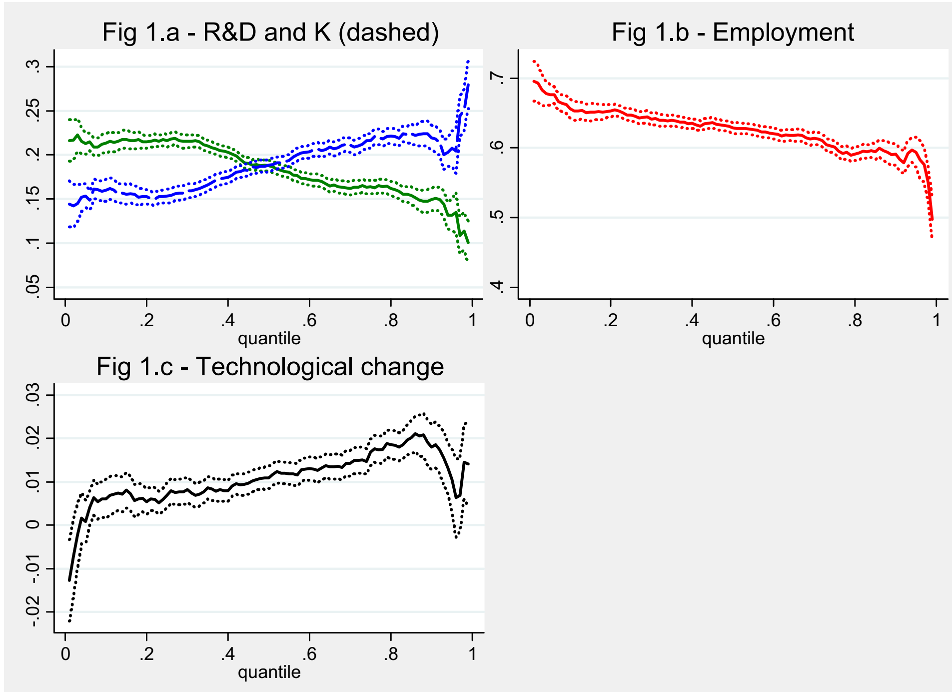

Fig. 1. Parameters' distribution from quantile regression – all sample.

estimates are not a reliable account of what happens along the firms' size distribution (in terms of turnover). This latter set of estimates – according to which, for the firms under investigation, output increases to a larger extent with the increase of their physical rather than knowledge capital[17] – is confirmed only by the largest firms of the whole sample (Fig. 1a). At the median quantile, the difference in the coefficients is not statistically significant. Moreover, the opposite result holds true for the first half of the size distribution, where the returns to physical capital are lower than those of knowledge capital. The increasing (decreasing) impact that physical (knowledge) capital shows along the distribution completes what can be deemed an expected picture.

The smallest innovative firms of the whole sample do not show relatively high returns from the exploitation of their physical capital. Conversely, investing in R&D from relatively lower scales of production is associated to a greater economic impact for them (Fig. 1a). The opposite can be said for the larger firms of the sample. Increasing the scale of their plants and machinery turns out to be relatively more productive than investing more in R&D. This is another interesting result of our quantile analysis. Although all the firms in the sample are large R&D investors, the economic exploitation of their R&D seems more connected with the innovative mode of the smaller of them: somehow surprisingly, larger innovative projects are associated to larger economic returns for the smaller firms of the sample. Once again, however, the quantile estimates per group of sectors introduce important specifications in this last respect (Fig. 2).

In mid-high (Fig. 2b) and low/mid-low tech sectors (Fig. 2c), the results of the average-based estimators seem to be confirmed along the quantiles: the output elasticity of physical capital is higher than that of knowledge capital, and this is also true, though to a lesser extent, for the smallest companies of the relative distribution (that is, its first quantiles). This might be explained by the technological regime of these sectors – in some way traceable to scale-intensive (mid-high) and supplier-dominated (low/mid-low) sectors – and by their intensity of physical capital. Furthermore, we should consider that, as the sample descriptive statistics show, the firms in these two groups of sectors are of larger size on average and could thus be better equipped for dealing with the indivisibility of physical capital investments. This is particularly evident in mid-high tech sectors (Fig. 2b), where the output elasticity of K becomes increasingly higher for larger quantiles of firms. At the same time, consistently with the results from the whole sample, the returns to R&D decrease with firm size in both sectors.

However, in the high-tech sectors (Fig. 2a) – and in this case only – the contribution of knowledge capital is larger than that of physical capital along the whole size distribution of the sample. For these firms, the sectoral pattern of innovation is such that R&D-based technological knowledge is the key factor in terms of production, irrespective of firm size; in other words, in these sectors, the different outcomes that we found associated to small and large firms in the aggregate sample do not appear to be as important. R&D is the pivotal input in high-tech, for whatever model of firm-size. All in all, this is another interesting, if expected, result, which supports other evidence on corporate R&D investments in high-tech sectors in Europe (Ortega-Argilés et al., 2010; Moncada-Paternò-Castello, 2011).

A final related bit of evidence on the marginal returns of the different capital inputs that firms use in production concerns their comparison across sectors. Quite interestingly, within the same quantile, the coefficients of knowledge capital are systematically higher in high-tech sectors than in mid-high and low/mid-low ones (Table 2). Conversely, within the same quantile, physical capital is correlated with output with systematically higher coefficients moving from mid-high to low/mid-low tech sectors. This result is also consistent with previous empirical work on the same set of Scoreboard companies (Ortega-Argilés et al., 2010, 2011), which suggests that, even for this homogeneous set of firms, productivity

[17] The elasticity of output with respect to physical and knowledge capital calculated with OLS is 20% and 17%, respectively. The RE estimates further exacerbate this difference, whereas the SYS-GMM seems to somehow overestimate the coefficient attached to the physical capital and underestimate that of labour (see Table 1).

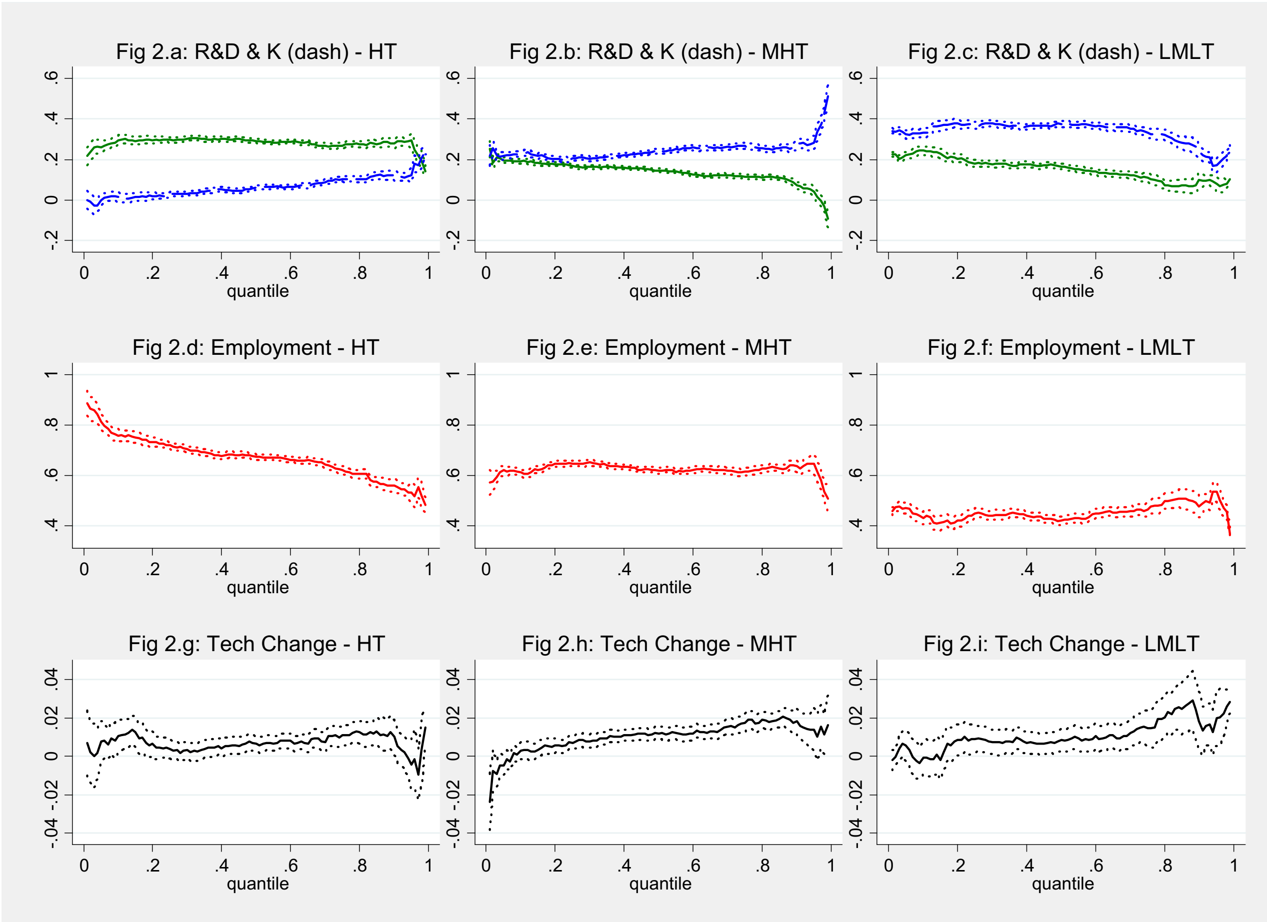

Fig. 2. Parameters' distribution from quantile regression – by technological sector.

in these different sectors is associated to a disembodied and embodied kind of technological change, respectively: a point that we will elaborate on in the following.

The previous results on the output elasticity of *K* and *RD* are even more interesting if we link them with size and sector variations in the economic impact of labour (*L*). Although we can only claim this in terms of associations, the whole sample of firms appears subject to an expected increase in mechanisation/automation in line with increased firm size (Todd and Oi, 1999): in terms of output impact, labour appears to be substituted by physical capital along the corresponding distribution (Fig. 1b). Sectoral estimations provide a more accurate interpretation of this result. Larger firms become progressively less associated to the pivotal economic impact of labour only in the high-tech sectors (Fig. 2d). Their stage of technological development and their relatively smaller average size might actually make a (physical) capitalisation process still relevant. In contrast, in the mid-high (Fig. 2e) and low/mid-low tech sectors (Fig. 2f), the technological regimes appear to be so mature and physical capital so intensive that the economic impact of labour remains constant along their size distribution.[18] This is more so for mid-high than low/mid-low tech sectors, whose output elasticity of labour is only about 2/3 of the former.

[18] To be sure, while the hypothesis that the within-sectors/across-quantiles coefficients of *K* and *L* are equal is never rejected at a 5% significance level in MHT and LMLT sectors; in the case of HT, the equality of the parameters estimated at the first, second and third quartile is always rejected.

To summarise, the analysis of the marginal returns of production factors shows important sector specificities in their use/impact for the firms under investigation. With the exception of the high-tech sector, for the top R&D spenders (and thus presumably innovative firms) in our sample, a greater production impact from an increase in size is not significantly correlated with a shift from physical to knowledge capital. The sectoral system of innovation appears more binding in this last respect.

Finally, we address the rate of technical progress that the estimate of the production function enables us to detect. A first interesting insight here comes from the quantile estimates for the whole sample of top R&D investors (Fig. 1c). Although they all rely heavily on R&D investments (at least in absolute terms) for their innovation activities, the increase of technological knowledge from which they benefit over time varies with their size: the larger the R&D investor, the higher its rate of technical progress. Once linked with the (similar) size variation that we have found along the whole sample for the marginal return of physical capital (Fig. 1a), this result would suggest an important tentative conclusion. For the firms that we are investigating, the most appreciable kind of technical progress seems to be of an embodied nature. In other words, at least without distinguishing by their economic sector of activity, the technical change of our top R&D investors becomes appreciable, provided it is associated with ameliorated plants and machinery for their production process.

This tentative result is, however, only partially confirmed by the quantile estimates at the sector level. In the mid-high tech sectors (Fig. 2h), and less univocally in low/mid-low ones (Fig. 2i), where we also find evidence of a larger relative impact of physical than knowledge capital along the entire size distribution, technical progress increases significantly with firm size, as it is at the aggregated level. In the high-tech sectors, on the other hand, where we previously found unique evidence of a general dominant impact of knowledge over physical capital, the rate of technical progress increases more smoothly over the relative size distribution (Fig. 2g).[19]

On the basis of these last results, we can more accurately state that the technological progress of the investigated firms appears embodied, and linked to the advantages that the larger companies in the sample have with respect to the smaller ones in investing in the expansion of their physical capital, but only in scale-intensive and supply dominated sectors. In high-tech sectors, by contrast, the size of the firms' plants does not seem to substantially interfere with their rate of technical change. In these sectors, where the economic impact of knowledge capital appears systematically larger than that of physical capital, and the average size is comparatively smaller, the hypothesis of a disembodied kind of technical change seems to be more plausible.

4.1. Robustness checks

A number of checks have been implemented in order to ascertain the robustness of our results. A first set of checks refers to the estimation of Eq. (3) by making use of the same pooled sample described in Section 3.1. First of all, in the construction of knowledge and capital stocks according to the Perpetual Inventory Method, we have made the rate of depreciation, δ, vary across our three groups of sectors, allowing for faster obsolescence in R&D and in the high-tech sectors. Accordingly, the rate δ for knowledge (physical) capital has been set to 15 (10), 12 (8), and 10 (6), for high-tech, medium/high-tech, and medium/low-tech, respectively. The relative results[20] hold true irrespectively from these alternatives. In addition, we have tried to fully exploit firm-level heterogeneity by calculating the average growth rate of *K* and *RD* at the firm level. However, the values of the average growth rates (in real terms) of one of the two types of investments turned out to be negative for some companies; for these companies we have thus not been able to calculate initial stock values. In order to avoid mixing different calculation strategies, we decided to take the average sectoral growth rates for all companies.

Finally, in order to control for the possibility of spurious correlation between employment (l) and output (y), in spite of the limited accuracy with which we have been able to capture it, *Age* has been inserted as a further firm-specific control in the estimate of Eq. (3).[21] Despite some minor changes, results[22] are consistent overall and lead us to exclude the risk of spurious correlations.

An additional set of robustness checks has been implemented by trying to retain the longitudinal dimension of the data, which spans over nine years of time (2002–2010). A first methodology for doing this would be the use of panel quantile regression, on which important progress has been recently made in theoretical econometrics (e.g. Koenker, 2004; Lamarche, 2010; Canay, 2011). In spite of the attention this methodology has been attracting, the number of published papers examining it is still limited. Furthermore, its empirical estimation is still hesitant and hampered by a number of methodological problems, in addition to the still burgeoning development of relative software routines (mainly in R). Among these problems, the estimation of the fixed effects is crucial, in order to account for "unobserved heterogeneity", which is not adequately controlled for by other covariates in the model. In this last respect, Koenker (2004) points out that, when the number of individual observations is relatively modest "it is quite unrealistic to attempt to estimate a τ-dependent, distributional, individual effect. At best we may be able to estimate an individual specific location-shift effect, and even this may strain credulity" (Koenker, 2004, p. 76). In a quantile framework, a large number of fixed effects can inflate the variability of the estimates of other covariates.[23] Koenker demonstrates that some degree of regularisation is desirable, but also points out that deciding precisely how much shrinkage should be imposed is quite challenging.[24]

More recently, Canay (2011) has proposed a simple estimator to deal with fixed effects, amounting to a pure location shifter, which is consistent and asymptotically normal as $T \rightarrow \infty$. However, some of the assumptions of the relative identification strategy do not fit with our exercise. In particular, fixed effects and unobservables are assumed independent one from each other, and the fixed effect does not change across the quantiles. Given the relatively short-time span of our panel (9 years), and the high mobility of companies across the quantiles over time (Google and HTC being two noteworthy examples among many others), these hypotheses crucially undermine the reliability of Canay's approach in our context.

Because of these difficulties, we opted for a sort of 'second-best' solution[25] and, in Table 3, we tested for the stability of the parameters of Eq. (3) estimated with respect to the pooled sample, by running yearly quantile estimations. Of course, given the lower number of degrees of freedom and the cross-sectional specification of the model year per year, the results are not expected to perfectly replicate those of the whole sample. Still, some alignment between the relative results should be desirable.

All in all, results are robust both within and across the three quantiles. In the first respect, the rate of change between the marginal returns of the three production factors for the pooled sample (last column), and those for the specific years, in general (with the exception of the first year of the series) does not overcome fifteen percentage points, being in many cases lower than ten points. Across the quantiles, labour is confirmed as the more productive factor for each year. More importantly, in the first quantile, knowledge capital overcomes the returns of physical capital all along the series, supporting our interpretation about the importance of a relatively smaller-size model of exploiting R&D in production. In the largest quantile (with the only exception of 2009), physical capital appears to be the driver of an embodied kind of technological change for each and every year, as we suggested above. In the intermediate quantile, as in the pooled sample, differences in returns are misty. With the support of a

[19] This is also confirmed by a series of Wald tests comparing the coefficients along the size distribution. The null hypothesis of equality of the parameters estimated at the first, second and third quartiles is never rejected.

[20] Available from the authors upon request.

[21] This control is motivated by the increasing empirical evidence of a relationship between age and economic (and innovative) output (Coad et al., 2013; Huergo and Jaumandreu, 2004), and that of a factual correlation between younger (older) and smaller (larger) firms. Unfortunately, the only available information in this last respect is the incorporation year of our companies, which, in many cases, is not accurate for calculating their age, because of possible mergers and acquisitions.

[22] Again available on request.

[23] The main issue with quantile panel data models is that the standard demeaning (or differencing) techniques used in standard models do not represent feasible approaches. When the number of panels is large, this entails a huge number of estimated α_i (individual effects) parameters. The so-called "regularisation" or "shrinkage" of these individual effects towards a common value could help in reducing this inflation effect.

[24] As confirmation of this issue, the panel quantile estimates that we have tried by using the 'rqpd' package developed for *R*, actually show large variations with small changes in the shrinkage parameter and are very sensible to the introduction/exclusion of our independent variables (in particular, the time trend and the different sets of dummy variables used).

[25] We are grateful to one of the reviewers for this suggestion.

Table 3
Comparison of the yearly quantile estimates coefficients.

	2002	2003	2004	2005	2006	2007	2008	2009	2010	All years
First quartile										
Knowledge capital	0.175***	0.197***	0.218***	0.228***	0.222***	0.234***	0.237***	0.227***	0.237***	***0.216***
	(5.889)	(8.149)	(8.721)	(8.715)	(15.880)	(14.191)	(10.037)	(10.747)	(8.032)	
Physical capital	0.153***	0.143***	0.149***	0.154***	0.160***	0.143***	0.132***	0.144***	0.174***	***0.153***
	(5.416)	(6.041)	(5.708)	(5.412)	(7.212)	(6.908)	(6.190)	(3.460)	(5.433)	
Employment	0.701***	0.675***	0.638***	0.634***	0.629***	0.645***	0.644***	0.649***	0.622***	***0.648***
	(24.167)	(19.560)	(18.681)	(22.872)	(19.681)	(28.606)	(23.021)	(17.533)	(18.158)	
Pseudo R-squared	*0.817*	*0.810*	*0.795*	*0.800*	*0.790*	*0.795*	*0.792*	*0.789*	*0.785*	
Second quartile										
Knowledge capital	0.152***	0.173***	0.193***	0.212***	0.185***	0.196***	0.211***	0.207***	0.201***	***0.188***
	(5.406)	(6.570)	(6.695)	(6.579)	(6.861)	(5.410)	(10.366)	(8.468)	(10.426)	
Physical capital	0.177***	0.205***	0.197***	0.219***	0.193***	0.188***	0.188***	0.166***	0.165***	***0.188***
	(7.875)	(7.066)	(6.262)	(9.831)	(6.670)	(7.251)	(6.332)	(6.074)	(5.336)	
Employment	0.674***	0.619***	0.605***	0.572***	0.627***	0.628***	0.614***	0.642***	0.646***	***0.629***
	(36.084)	(22.136)	(22.952)	(17.353)	(22.123)	(23.009)	(21.939)	(24.113)	(16.645)	
Pseudo R-squared	*0.811*	*0.806*	*0.794*	*0.796*	*0.790*	*0.790*	*0.788*	*0.789*	*0.788*	
Third quartile										
Knowledge capital	0.117***	0.159***	0.165***	0.170***	0.173***	0.183***	0.187***	0.197***	0.171***	***0.164***
	(2.936)	(4.100)	(5.776)	(5.479)	(5.336)	(7.263)	(7.100)	(5.079)	(4.994)	
Physical capital	0.221***	0.223***	0.252***	0.223***	0.228***	0.191***	0.197***	0.143***	0.190***	***0.216***
	(8.538)	(10.608)	(11.175)	(10.765)	(7.391)	(5.394)	(5.387)	(4.786)	(6.108)	
Employment	0.645***	0.600***	0.560***	0.577***	0.575***	0.616***	0.617***	0.662***	0.636***	***0.602***
	(14.391)	(17.037)	(15.220)	(17.930)	(14.292)	(14.875)	(16.576)	(17.314)	(21.076)	
Pseudo R-squared	*0.801*	*0.795*	*0.785*	*0.785*	*0.780*	*0.777*	*0.775*	*0.779*	*0.780*	

Bold and italics has been used to highlight the quartile of interest (it is a kind of internal title).
t-Statistics from bootstrapped regressions (50 replications). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

quite large pseudo *R*-squared, these estimates appear aligned with the main ones we reported in the previous section and increase the validity of the results that we have identified.[26]

As a final robustness check, following Hauk and Wacziarg (2009), we have controlled for possible measurement errors by applying OLS and quantile estimations on the averaged values of the variables considered. In this case as well, results are robust and confirm those of our benchmark model.[27]

5. Conclusions

Top R&D investors are inherently diverse, not only in the realm of innovation but also in that of production. The quantile estimation of their production function – augmented for the role of knowledge capital – reveals important elements of heterogeneity that standard estimations would otherwise hide. In particular, their size intertwines with their economic sector in specifying some basic production-related issues, which would otherwise be considered of a general nature for the investigated firms.

This result has important methodological implications for research on the issue. While the use of quantile estimates is becoming increasingly popular for detecting firm-specific factors, our application suggests that the attention given to heterogeneity deriving from firm size should not be viewed in isolation of that originating from the sector in which any given firm operates. Furthermore, our results suggest that technical efficiency measures could be biased when the underlying heterogeneity in the input factors is not taken into account.

Our results also have some interesting policy implications. First of all, although they are all quite large R&D spenders, the extent to which the innovative firms in our sample benefit from returns to scale is remarkable. Returns to scale appear to be decreasing only for the largest companies in the sample, which are mainly located in the lower-tech sectors. In high-tech sectors, on the other hand, returns to scale in production still appear exploitable. This is of great relevance when we think about policy support for the growth of innovative companies (in our case, innovative investors). While such a stimulus is usually considered suitable mainly for small (and new) technology-based firms, our evidence suggests that larger firms could also benefit from it, since they are not constrained by problems of efficiency in production.

Sector-specific effects are also important when we look at the production impact of the different inputs that firms employ. The output of our companies correlates substantially with changes in their knowledge capital only in the case of high-tech sectors. Conversely, in lower-tech sectors, where firm size is relatively larger on average, physical capital appears to be the pivotal production input along the whole of firm size distribution. This is an interesting result when we look at the recent literature (mainly at the country-sector level) about the impact of tangible vs. intangible assets (e.g. Corrado et al., 2009). By referring to our sample of top R&D spenders, tangible assets appear to count substantially more than intangible ones, unless we refer to the firms of smaller size and higher technological level in our sample, which are the only ones that actually appear "knowledge intensive". Furthermore, the policy implication of this result is quite important and somehow in line with that obtained by other studies on the same sample of top R&D spenders, which instead focus on their labour productivity (Kumbhakar et al., 2012). Policy support to R&D would have the greatest impact (economic, in our case) in high-tech sectors, whereas the other economic sectors would benefit more from incentives and/or fiscal facilities to physical capital investments. Also when looking at the production realm, policies for innovative firms need to be tailored.

Related to this result is the one we obtained for the production impact of labour across the three groups of sectors that we

[26] Similar checks have been carried out with respect to the estimates by sector, with similarly consistent results (available on request).

[27] Hauk and Wacziarg (2009) show that, in the presence of measurement errors, an OLS applied to a single cross-section of variables averaged over time (between estimator) outperforms other estimators, such as the fixed-effect or the GMM. In particular, in terms of average absolute bias, it reduces the variance of the measurement error relative to the true signal. Also for this last check, whose results are available upon request, we are grateful to one of the reviewers.

consider. In mid-high tech sectors, this is on average lower than in high-tech sectors. However, an important distinction appears between the two along their respective size distributions. In the high-tech sector, while the output elasticity of knowledge capital is size invariant, that of elasticity of labour decreases with firm size, hinting at its substitution by physical capital. This is consistent with a progressively higher degree of automation with increased firm size. In mid-high tech sectors, by contrast, the economic importance of labour remains invariant along the size distribution; the same holds true in the low/mid-low tech sectors, although at a lower average level. As we have said, what is noticeable here is rather a size-dependent substitution effect of knowledge for physical capital. On this basis, an interesting policy implication could accompany those we have provided above, concerning the opportunity of supporting physical capital investments in the lower-tech sectors. Because of the maturity stage of the relative technology, this policy support is unlikely to generate labour substitution effects: employment is expected to keep its relevance, independently of firm size.

The need to tailor support to R&D investors on the basis of relevant production inputs also emerges from the technical progress that our approach enables us to detect. In this regard, the results we obtain are most connected to the innovative performance of our firms and to the innovative policies which can act on it. In the mid-high and low/mid-low sectors, our estimates provide evidence of technological change of an embodied nature, for which high intensity of physical capital and large company size provide an important advantage. Conversely, in high-tech sectors, opportunities for technical change appear to be of a more disembodied kind, with no advantages for larger firms with larger capital stocks. This last result holds true in the presence of the dominant role of knowledge capital over physical capital, along the entire size distribution. Taking into account the specificities that technical change reveals in different sectors with respect to its embodied and disembodied nature, the need for a sector focus regarding R&D policies is thus confirmed.

Acknowledgements

Previous versions of this paper were presented at the 35th DRUID Celebration Conference (June 17–19, 2013, Barcelona – Spain), the KfG/IS inter-Service Scientific Seminars Series at JRC (June 4, 2013, Seville – Spain), and the Enabling Global Transitions through Innovation Conference (May 8–9, 2013, Istanbul – Turkey). We are grateful to the participants of these events for their comments and suggestions. Furthermore, we are indebted to Giuseppe Vittucci Marzetti and to three anonymous referees for their useful insights.

Disclaimer: The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

Appendix.

See Tables A1 and A2.

Table A1
Descriptive statistics of the sample.

	All sample	High-tech	Medium/high tech	Medium/low tech
No. of observations	8,990	3,621	3,773	1,596
Net sales (mil. €)				
Average	8,088	3,727	8,266	17,560
Standard deviation	19,373	10,259	18,343	30,784
Median	1,846	635	2,553	7,694
R&D Investments (mil. €)				
Average	304	364	311	152
Standard deviation	777	877	807	295
Median	67	70	67	63
Capital expenditure (mil. €)				
Average	552	194	492	1,508
Standard deviation	1,720	583	1,692	2,815
Median	73	24	96	443
Employment (# of emp.)				
Average	28,016	14,835	32,048	48,387
Standard deviation	55,686	38,942	54,910	77,820
Median	8,336	3,034	11,821	21,742

Table A2
Industry classification by sector groups.[a]

Sector groups	Industries
High tech sectors (R&D intensity above 5%)	Pharmaceuticals & biotechnology; Health care equipment & services; Technology hardware & equipment; Software & computer services
Medium/high tech sectors (R&D intensity between 2% and 5%)	Electronics & electrical equipment; Automobiles & parts; Aerospace & defence; Industrial engineering & machinery; Chemicals; Personal goods; Household goods; General industrials; Support services
Medium/low tech sectors (R&D intensity below 2%)	Food producers; Beverages; Travel & leisure; Media; Oil equipment; Electricity; Fixed line telecommunications; Oil & gas producers; Industrial metals; Construction & materials; Food & drug retailers; Transportation; Mining; Tobacco; Multi-utilities

[a] IRI scoreboard sector groups by R&D intensity; ICB (Industry Classification Benchmark), 4-digit level.

References

Acs, Z.J., Audretsch, D.B., 1987. Innovation in large and small firms. Economics Letters 23 (1), 109–112.

Aghion, P., Howitt, P., 1992. A model of growth through creative destruction. Econometrica 60 (2), 323–351.

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo. Evidence and an application to employment equations. Review of Economic Studies 58, 277–297.

Bartelsman, E., Dobbelaere, S., Peters, B., 2013. Allocation of human capital and innovation at the frontier: firm-level evidence on Germany and the Netherlands. IZA DP No. 7540.

Battese, G.E., Broca, S.S., 1997. Functional forms of stochastic frontier production functions and models for technical inefficiency effects: a comparative study for wheat farmers in Pakistan. Journal of Productivity Analysis 8 (4), 395–414.

Bernini, C., Freo, M., Gardini, A., 2004. Quantile estimation of frontier production function. Empirical Economics 29, 373–381.

Blundell, R.W., Bond, S.R., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115–143.

Blundell, R.W., Bond, S.R., 2000. GMM estimation with persistent panel data: an application to production functions. Econometric Reviews 19 (3), 321–340.

Breschi, S., Malerba, F., Orsenigo, L., 2000. Technological regimes and Schumpeterian patterns of innovation. Economic Journal 110 (463), 388–410.

Brown, M., Popkin, J., 1962. A measure of technological change and returns to scale. Review of Economics and Statistics 44 (4), 402–411.

Canay, I.A., 2011. A simple approach to quantile regression for panel data. Econometrics Journal 14, 368–386.

Castellacci, F., 2008. Technological paradigms, regimes and trajectories: manufacturing and service industries in a new taxonomy of sectoral patterns of innovation. Research Policy 37 (6), 978–994.

Cincera, M., Ravet, J., 2010. Financing constraints and R&D investments of large corporations in Europe and the US. Science and Public Policy 37 (6), 455–466.

Ciriaci, D., Moncada-Paternò-Castello, P., Voigt, P., 2012. Does size or age of innovative firms affect their growth persistence? Evidence from a panel of innovative Spanish firms. IPTS Working Papers on Corporate R&D and Innovation, 3/2012.

Coad, A., Segarra, A., Teruel, M., 2013. Like milk or wine: does firm performance improve with age? Structural Change and Economic Dynamics 24, 173–189.

Coad, A., Rao, R., 2006. Innovation and market value: a quantile regression analysis. Economics Bulletin 15 (13), 1–10.

Coad, A., Rao, R., 2008. Innovation and firm growth in high-tech sectors: a quantile regression approach. Research Policy 37 (4), 633–648.

Cobbold, T., 2003. A Comparison of Gross Output and Value-Added Methods of Productivity Estimation. Productivity Commission Research Memorandum, Canberra.

Cohen, W.M., 2010. Fifty years of empirical studies of innovative activity and performance. Handbook of the Economics of Innovation, vol. 1., pp. 129–213.

Cohen, A.J., Harcourt, G.C., 2003. Whatever happened to the Cambridge capital theory controversies? Journal of Economic Perspectives 17 (1), 99–214.

Cohen, W.M., Klepper, S., 1996. Firm size and the nature of innovation within industries: the case of process and product R&D. Review of Economics and Statistics 78 (2), 232–243.

Conte, A., Vivarelli, M., 2005. One or many knowledge production functions? Mapping innovative activity using microdata. IZA Discussion Papers 1878. Institute for the Study of Labor (IZA).

Corrado, C., Hulten, C., Sichel, D., 2009. Intangible capital and US economic growth. Review of Income and Wealth 55 (3), 661–685.

Crepon, B., Duguet, E., Mairesse, J., 1998. Research, innovation and productivity: an econometric analysis at the firm level. Economics of Innovation and New Technology 7 (2), 115–158.

Czarnitzki, D., Kraft, K., Thorwarth, S., 2009. The knowledge production of 'R' and 'D'. Economics Letters 105 (1), 141–143.

D'Agostino, L.M., Laursen, K., Santangelo, G.D., 2013. The impact of R&D offshoring on the home knowledge production of OECD investing regions. Journal of Economic Geography 13 (1), 145–175.

Dosi, G., 1988. Sources, procedures, and microeconomic effects of innovation. Journal of Economic Literature, 1120–1130.

Douglas, P.H., 1976. The Cobb–Douglas production function once again: its history, its testing, and some new empirical values. Journal of Political Economy 84, 903–915.

Ebersberger, B., Marsili, O., Reichstein, T., Salter, A., 2010. Into thin air: using a quantile regression approach to explore the relationship between R&D and innovation. International Review of Applied Economics 24 (1), 95–102.

Evangelista, R., Vezzani, A., 2012. The impact of technological and organizational innovations on employment in European Firms. Industrial and Corporate Change 21 (4), 871–899.

Evangelista, R., Vezzani, A., 2010. The economic impact of technological and organizational innovations. A firm-level analysis. Research Policy 39, 1253–1260.

Falk, M., 2012. Quantile estimates of the impact of R&D intensity on firm performance. Small Business Economics 39 (1), 19–37.

Foss, N.J., 1994. The theory of the firm: the Austrians as precursors and critics of contemporary theory. Review of Austrian Economics 7 (1), 31–65.

García-Manjóna, J.V., Romero-Merino, M.E., 2012. Research, development, and firm growth. Empirical evidence from European top R&D spending firms. Research Policy 41 (6), 1084–1090.

Gordon, R.J., 1990. The Measurement of Durable Goods Prices. University of Chicago Press, Chicago and London.

Greenwood, J., Hercowitz, Z., Krusell, P., 1997. Long run implications of investment specific technological change. American Economic Review 87 (3), 342–362.

Griliches, Z., 1998. The search for R&D spillovers, NBER chapters. In: R&D and Productivity: The Econometric Evidence. National Bureau of Economic Research, pp. 251–268.

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

Griliches, Z., Mairesse, J., 1995. Production Functions: The Search for Identification. National Bureau of Economic Research No. 5067.

Hall, B.H., Mairesse, J., 1995. Exploring the relationship between R&D and productivity in French manufacturing firms. Journal of Econometrics 65 (1), 263–293.

Hauk, W., Wacziarg, R., 2009. A Monte Carlo study of growth regressions. Journal of Economic Growth 14 (2), 103–147.

Heidenreich, M., 2009. Innovation patterns and location of European low- and medium-technology industries. Research Policy 38, 483–494.

Huergo, E., Jaumandreu, J., 2004. Firms' age, process innovation and productivity growth. International Journal of Industrial Organization 22 (4), 541–559.

IRI-IPTS, European Commission, 2013. EU R&D Scoreboard. The 2013EU Industrial R&D Investment Scoreboard. Publications Office of the European Union, Luxembourg.

Jorgenson, D.W., 1966. The embodiment hypothesis. Journal of Political Economy 74 (1), 1–17.

Kaiser, U., 2009. Patents and profit rates. Economics Letters 104 (2), 79–80.

Kamien, M., Schwartz, N., 1982. Market Structure and Innovation. Cambridge University Press, Cambridge.

Koenker, R., 2004. Quantile regression for longitudinal data. Journal of Multivariate Analysis 91 (1), 74–89.

Koenker, R., Hallock, K.F., 2001. Quantile regression. Journal of Economic Perspectives 15 (4), 143–156.

Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46 (1), 33–50.

Kumbhakar, S., Ortega-Argilés, R., Potters, L., Vivarelli, M., Voigt, P., 2012. Corporate R&D and firm efficiency: evidence from Europe's top R&D investors. Journal of Productivity Analysis 37 (2), 125–140.

Lamarche, C., 2010. Robust penalized quantile regression estimation for panel data. Journal of Econometrics 157, 396–408.

Lööf, H., Heshmati, A., 2002. Knowledge capital and performance heterogeneity: a firm-level innovation study. International Journal of Production Economics 76 (1), 61–85.

Mairesse, J., Jaumandreu, J., 2005. Panel-data estimates of the production function and the revenue function: what difference does it make? Scandinavian Journal of Economics 107 (4), 651–672.

Malerba, F., 2002. Sectoral systems of innovation and production. Research Policy 31 (2), 247–264.

Malerba, F., Orsenigo, L., 1996. Schumpeterian patterns of innovation are technology-specific. Research Policy 25, 451–478.

Malerba, F., Orsenigo, L., 1995. Schumpeterian patterns of innovation. Cambridge Journal of Economics 19, 47–65.

Mata, J., Wörter, M., 2013. Risky innovation: the impact of internal and external R&D strategies upon the distribution of returns. Research Policy 42 (2), 495–501.

Moncada-Paternò-Castello, P., 2011. Companies' growth in the EU: what is research and innovation policy's role? IPTS Working Paper on Corporate R&D and Innovation, 3/2011.

Montobbio, F., 2003. Sectoral patterns of technological activity and export market share dynamics. Cambridge Journal of Economics 27, 523–545.

Montobbio, F., 2002. An evolutionary model of industrial growth and structural change. Structural Change and Economic Dynamics 13 (4), 387–414.

Nahm, J.W., 2001. Non-parametric quantile regression analysis of R&D-sales relationship for Korean firms. Empirical Economics 26 (1), 259–270.

Olley, S., Pakes, A., 1996. The dynamics of productivity in the telecommunications equipment industry. Econometrica 64, 1263–1270.

Ortega-Argilés, R., Potters, L., Vivarelli, M., 2011. R&D and productivity: testing sectoral peculiarities using micro data. Empirical Economics 41, 817–839.

Ortega-Argilés, R., Piva, M., Potters, L., Vivarelli, M., 2010. Is corporate R&D investment in high-tech sectors more effective? Contemporary Economic Policy 28, 353–365.

Pakes, A., Griliches, Z., 1984. Patents and R&D at the firm level: a first look. In: Griliches, Z. (Ed.), R&D, Patents, and Productivity. University of Chicago Press, pp. 55–72.

Pavitt, K., 1984. Sectoral patterns of technical change: towards a taxonomy and a theory. Research Policy 13 (6), 343–373.

Ponds, R., Van Oort, F., Frenken, K., 2010. Innovation, spillovers and university-industry collaboration: an extended knowledge production function approach. Journal of Economic Geography 10 (2), 231–255.

Ramani, S.V., El-Aroui Carrére, M.-A., Carrére, M., 2008. On estimating a knowledge production function at the firm and sector level using patent statistics. Research Policy 37 (9), 1568–1570.

Romer, P.M., 1990. Endogenous technological change. Journal of Political Economy 98 (5), S71–S102.

Rosenberg, N., 1982. Inside the Black Box: Technology and Economics. Cambridge University Press, Cambridge.

Rosenberg, N., 1994. Exploring the Black Box: Technology, Economics, and History. Cambridge University Press, Cambridge.

Salter, W.E.G., 1960. Productivity and Technical Change. Cambridge University Press, Cambridge.

Santamaría, L., Nieto, M.J., Barge-Gil, A., 2009. Beyond formal R&D: taking advantage of other sources of innovation in low- and medium-technology industries. Research Policy 38 (3), 507–517.

Scherer, F.M., 1965. Firm size, market structure, opportunity, and the output of patented inventions. American Economic Review 55 (5), 1097–1100.

Segarra, A., Teruel, M., 2011. Productivity and R&D sources: evidence for Catalan firms. Economics of Innovation and New Technology 20 (8), 727–748.

Shephard, R.W., 1974. The econometrician's viewpoint of a production function. Zeitschrift fiir NationalSkonomie 34, 403–408.

Solow, R.M., 1960. Investment and technological progress. In: Arrow, K.J., Karlin, S., Suppes, P. (Eds.), Mathematical Methods in the Social Sciences 1959. Stanford University Press, Stanford, CA, pp. 89–104.

Stam, E., Wennberg, K., 2009. The roles of R&D in new firm growth. Small Business Economics 33 (1), 77–89.

Todd, L.I., Oi, W.Y., 1999. Workers are more productive in large firms. American Economic Review 89 (2), 104–108, Papers and Proceedings of the One Hundred Eleventh Annual Meeting of the American Economic Association.

Tone, K., Sahoo, B.K., 2003. Scale, indivisibilities and production function in data envelopment analysis. International Journal of Production Economics 42 (2), 165–192.

Utterback, J.M., Abernathy, W.J., 1975. A dynamic model of process and product innovation. Omega 3 (6), 639–656.

Yorukoglu, M., 1998. The information technology productivity paradox. Review of Economic Dynamics 1 (2), 551–592.

Zucker, L.G., Darby, M.R., Furner, J., Liu, R.C., Ma, H., 2007. Minerva unbound: knowledge stocks, knowledge flows and new knowledge production. Research Policy 36 (6), 850–863.