R&D and Productivity: Estimating Endogenous Productivity

ULRICH DORASZELSKI

University of Pennsylvania

and

JORDI JAUMANDREU

Boston University

We develop a model of endogenous productivity change to examine the impact of the investment in knowledge on the productivity of firms. Our dynamic investment model extends the tradition of the knowledge capital model of Griliches (1979) that has remained a cornerstone of the productivity literature. Rather than constructing a stock of knowledge capital from a firm's observed R&D expenditures, we consider productivity to be unobservable to the econometrician. Our approach accounts for uncertainty, non-linearity, and heterogeneity across firms in the link between R&D and productivity. We also derive a novel estimator for production functions in this setting. Using an unbalanced panel of more than 1800 Spanish manufacturing firms in nine industries during the 1990s, we provide evidence of non-linearities as well as economically significant uncertainties in the R&D process. R&D expenditures play a key role in determining the differences in productivity across firms and the evolution of firm-level productivity over time.

Key words: R&D, Productivity, Knowledge capital model, Production function estimation

JEL Codes: D24, L60, O30

1. INTRODUCTION

A firm invests in R&D and related activities to develop and introduce product and process innovations. These investments in knowledge enhance the productivity of the firm and change its competitive position relative to that of other firms.

Our goal in this article is to assess the role of R&D in determining the differences in productivity across firms and the evolution of firm-level productivity over time. To achieve this goal, we develop a model of endogenous productivity change resulting from investment in knowledge. We also derive an estimator for production functions in this setting. In addition to the parameters of the production function, our estimator recovers the law of motion for productivity. With these tools in hand we study the link between R&D and productivity in Spanish manufacturing firms during the 1990s.

Our starting point is a dynamic model of a firm that invests in R&D in order to improve its productivity over time in addition to carrying out a series of investments in physical capital. Both investment decisions depend on the current productivity and the capital stock of the firm as do the subsequent decisions on static inputs such as labour and materials. Productivity follows a Markov process that can be shifted by R&D expenditures. The evolution of productivity is thus subject to random shocks. These innovations to productivity capture the factors that have a persistent influence on productivity such as absorption of techniques, modification of processes, uncertainties related to investments in physical capital, and gains and losses due to changes in labour composition and management abilities. For a firm that engages in R&D, the productivity innovations additionally capture the uncertainties inherent in the R&D process such as chance in discovery, degree of applicability, and success in implementation.

Our model of endogenous productivity change is not the first attempt to account for investment in knowledge. In a very influential paper, Griliches (1979) proposed to augment a standard production function with "a measure of the current state of technical knowledge, determined in part by current and past research and development expenditures" (p. 95). In practice, a firm's observed R&D expenditures are used to construct a proxy for the state of knowledge. This knowledge capital model has remained a cornerstone of the productivity literature for more than 25 years and has been applied in hundreds of studies on firm-level productivity (see the surveys by Mairesse and Sassenou (1991), Griliches (1995, 2000), and Hall *et al.* (2010)).

In a departure from the previous literature we do not attempt to construct a stock of knowledge capital from the available history of R&D expenditures and with it control for the impact of R&D on productivity. Instead, we consider productivity to be unobservable to the econometrician and in this way relax the assumptions on the R&D process in a natural fashion. Building on Hall and Hayashi's (1989) and Klette's (1996) extension of the knowledge capital model, we recognize that the outcome of the R&D process is likely to be subject to a high degree of uncertainty. Once discovered, an idea has to be developed and applied, and there are the technical and commercial uncertainties linked to its practical implementation. We further recognize, again building on Hall and Hayashi (1989) and Klette (1996), that current and past investments in knowledge are likely to interact with each other. Since there is little reason to believe that features such as complementarities and economies of scale in the accumulation of knowledge or the obsolescence of previously acquired knowledge can be adequately captured by simple functional forms, we deviate from Hall and Hayashi (1989) and Klette (1996) by modelling the interactions between current and past investments in knowledge in a flexible fashion.

To retrieve productivity at the level of the firm, we have to estimate the parameters of the production function. However, if a firm adjusts to a change in its productivity by expanding or contracting its production depending on whether the change is favourable or not, then unobserved productivity and input usage are correlated and biased estimates result (Marschak and Andrews, 1944). Recent advances in the structural estimation of production functions, starting with the dynamic investment model of Olley and Pakes (1996) (OP), tackle this endogeneity problem.[1] The insight of OP is that if observed investment is a monotone function of unobserved productivity, then this function can be inverted to back out—and thus control for—productivity. This line of research has been extended by Levinsohn and Petrin (2003) (LP) and Ackerberg *et al.* (2006) (ACF).

Instead of relying on the firm's dynamic programming problem as OP do, we use the fact that static inputs are decided on with current productivity known and therefore contain information about it. As first shown by LP, the input demands resulting from short-run profit maximization are

---

1. See Griliches and Mairesse (1998) and Ackerberg *et al.* (2007) for reviews of this and other problems that arise in the estimation of production functions.

invertible functions of unobserved productivity. We use this insight to control for productivity and obtain consistent estimates of the parameters of the production function. In addition, we recognize that, given a parametric specification of the production function, the functional form of the inverse input demand functions is known. As pointed out by Marschak and Andrews (1944), the structural assumptions imply parameter restrictions between the production function and the inverse input demand functions. Hence, we do not have to rely on non-parametric methods to estimate these functions. Because our parametric inversion fully exploits the parameter restrictions between the production function and the inverse input demand functions, it yields a particularly simple estimator for production functions.

We apply our estimator to an unbalanced panel of more than 1800 Spanish manufacturing firms during the 1990s. Our data are of notably high quality and combines information on production with information on firms' R&D activities in nine industries. This broad coverage of industries is uncommon and allows us to examine the link between R&D and productivity in a variety of settings that differ greatly in the importance of R&D. At the same time, it allows us to put to the test our model of endogenous productivity change and the estimator we develop for it.

Somewhat unusually we have firm-level wage and price data.[2] The fact that the wage and prices vary across firms is at variance with the often-made assumption in the literature following OP that all firms face the same wage and prices and that these variables can therefore be replaced by a dummy. Instead, as LP point out, the wage and prices must be accounted for in the inverse input demand functions used to recover unobserved productivity. While the non-parametric methods in OP, LP, and ACF extend to our model of endogenous productivity change, non-parametrically estimating the inverse input demand functions becomes more demanding on the data as the number of their arguments increases. Our parametric inversion circumvents this "curse of dimensionality" in non-parametric estimation and is much less demanding on the data.

Our estimates of the law of motion for productivity attest to important non-linearities and uncertainties in the R&D process. The impact of current R&D on future productivity depends crucially on current productivity. Non-linearities often take the form of complementarities between current R&D and current productivity. Furthermore, the R&D process must be treated as inherently uncertain. We estimate that, depending on the industry, between 25% and 75% of the variance in productivity is explained by innovations that cannot be predicted when decisions on R&D expenditures are made. Moreover, the return to R&D is often twice that of the return to investment in physical capital. Our estimates therefore suggest that the uncertainties inherent in the R&D process are economically significant and matter for firms' investment decisions.

Capturing the uncertainties in the R&D process also paves the way for heterogeneity across firms. Because we allow the shocks to productivity to accumulate over time, even firms with the same time path of R&D expenditures may not have the same productivity. This gives us the ability to assess the role of R&D in determining the differences in productivity across firms and the evolution of firm-level productivity over time.

Despite the uncertainties in the R&D process, we show that the expected productivity of firms that perform R&D is systematically more favourable in the sense that their distribution of expected productivity tends to stochastically dominate the distribution of firms that do not perform R&D. Assuming that the productivity process is exogenous (as in most of the literature following OP) amounts to averaging over firms with distinct innovative activities and hence blurs important differences in the impact of the investment in knowledge on the productivity of

---

2. There are other data sets such as the Colombian Annual Manufacturers Survey (Eslava *et al.*, 2004) and the Longitudinal Business Database at the U.S. Census Bureau that contain separate information about prices and quantities, at least for a subset of industries (Roberts and Supina, 1996; Foster *et al.*, 2008).

firms.[3] In addition, we estimate that firms that perform R&D contribute between 65% and 90% of productivity growth in the industries with intermediate or high innovative activity. Investments in R&D and related activities are thus a primary source of productivity growth.

Our model allows us to recover the entire distribution of the elasticity of output with respect to R&D expenditures—a measure of the return to R&D—as well as that of the elasticity of output with respect to already attained productivity—a measure of the degree of persistence in the productivity process. On average we obtain higher elasticities with respect to R&D expenditures than in the knowledge capital model and lower elasticities with respect to already attained productivity. Since old knowledge is hard to keep but new knowledge is easy to add, productivity is considerably more fluid than what the knowledge capital literature suggests.

The remainder of this article is organized as follows. Section 2 lays out a dynamic investment model with endogenous productivity change. Section 3 describes the data and Section 4 develops our empirical strategy. Sections 5 and 6 discuss our results and Section 7 concludes.

2. A MODEL FOR INVESTMENT IN KNOWLEDGE

A firm carries out two types of investments, one in physical capital and another in knowledge through R&D expenditures. Investment decisions are made in a discrete time setting with the goal of maximizing the expected net present value of future cash flows. Capital is the only dynamic (or "fixed") input among the conventional factors of production and accumulates according to $K_{jt} = (1-\delta)K_{jt-1} + I_{jt-1}$, where K_{jt} is the stock of capital of firm j in period t and δ is the rate of depreciation. This law of motion implies that investment I_{jt-1} chosen in period $t-1$ becomes productive in period t.

The productivity of firm j in period t is ω_{jt}. While the firm knows its productivity when it makes its decisions, we follow OP and often refer to ω_{jt} as "unobserved productivity" since it is not observed by the econometrician. Productivity is correlated over time and perhaps also correlated across firms. Because our goal is to assess the role of R&D in determining the differences in productivity across firms and the evolution of firm-level productivity over time, we have to endogenize the productivity process. To this end, we assume that productivity is governed by a controlled first-order Markov process with transition probabilities $P(\omega_{jt}|\omega_{jt-1}, r_{jt-1})$, where r_{jt-1} is the log of R&D expenditures.[4]

Adopting the convention that lower case letters and upper case letters denote logs and levels, respectively, the firm has the Cobb–Douglas production function

$$y_{jt} = \beta_0 + \beta_t t + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \omega_{jt} + e_{jt}, \tag{1}$$

where y_{jt} is the output of firm j in period t, l_{jt} is labour, and m_{jt} is materials. We allow for a secular trend t in the production function that we model as either a linear time trend or dummies (see Section 4.2 for details). In contrast to productivity ω_{jt}, the mean zero random shock e_{jt} is uncorrelated over time and across firms. The firm does not know the value of e_{jt} when it makes its decisions at time t.

The Bellman equation for the firm's dynamic programming problem is

$$V(s_{jt}) = \max_{i_{jt}, r_{jt}} \Pi(s_{jt}) - C_i(i_{jt}) - C_r(r_{jt}) + \frac{1}{1+\rho} E\big[V(s_{jt+1})|s_{jt}, i_{jt}, r_{jt}\big], \tag{2}$$

3. Subsequent research on extending the two-stage procedures in OP, LP, and ACF to an endogenous productivity process echoes this conclusion (De Loecker, 2010).

4. The literature following OP typically assumes an exogenous Markov process with transition probabilities $P(\omega_{jt}|\omega_{jt-1})$.

where $s_{jt} = (t, k_{jt}, \omega_{jt}, w_{jt}, p_{Mjt}, d_{jt})$ is the vector of state variables. Besides the trend t, the stock of capital k_{jt}, and productivity ω_{jt}, the vector of state variables comprises other variables that are correlated over time, namely the wage w_{jt} and the price of materials p_{Mjt} that the firm faces, and the demand shifter d_{jt}. The discount rate is ρ. Per-period profits are given by the indirect profit function $\Pi(\cdot)$ and the cost of investment in physical capital and knowledge by the cost functions $C_i(\cdot)$ and $C_r(\cdot)$, respectively. In the simplest case $C_i(\cdot)$ and $C_r(\cdot)$ just transform logs into levels, but they may also capture indivisibilities in investment projects or adjustment costs; their exact forms are irrelevant for our purposes.[5] In practice, investment opportunities and the price of equipment goods are likely to vary and the cost of investment in knowledge depends greatly on the nature of the undertaken project (Adda and Cooper, 2003, p. 188). To capture variation in the cost of investment in physical capital and knowledge across firms and time, a cost shifter x_{jt} can be added to $C_i(\cdot)$ and $C_r(\cdot)$ (and thus becomes part of s_{jt}). The dynamic programming problem gives rise to two policy functions $i(s_{jt})$ and $r(s_{jt})$ for the investments in physical capital and knowledge, respectively.

The firm anticipates the effect of R&D on productivity in period t when making the decision about investment in knowledge in period $t-1$. The Markovian assumption implies

$$\omega_{jt} = E\big[\omega_{jt}|\omega_{jt-1}, r_{jt-1}\big] + \xi_{jt} = g(\omega_{jt-1}, r_{jt-1}) + \xi_{jt}. \tag{3}$$

That is, *actual productivity* ω_{jt} in period t can be decomposed into *expected productivity* $g(\omega_{jt-1}, r_{jt-1})$ and a random shock ξ_{jt}. While the conditional expectation function $g(\cdot)$ depends on already attained productivity ω_{jt-1} and R&D expenditures r_{jt-1}, ξ_{jt} does not: by construction ξ_{jt} is mean independent (although not necessarily fully independent) of ω_{jt-1} and r_{jt-1}. This *productivity innovation* represents the uncertainties that are naturally linked to productivity plus the uncertainties inherent in the R&D process such as chance in discovery, degree of applicability, and success in implementation. It is important to stress the timing of decisions in this context: when the decision about investment in knowledge is made in period $t-1$, the firm is only able to anticipate the expected effect of R&D on productivity in period t as given by $g(\omega_{jt-1}, r_{jt-1})$ while its actual effect also depends on the realization of the productivity innovation ξ_{jt} that occurs after the investment has been completely carried out. The conditional expectation function $g(\cdot)$ is not observed by the econometrician and must be estimated non-parametrically along with the parameters of the production function.

As noted by LP, the firm's dynamic programming problem—and the policy functions for the investments in physical capital and knowledge it gives rise to—can easily become very complicated. Buettner (2005) incorporates R&D expenditures into the dynamic investment model of OP. To prove that the policy function for investment in physical capital can be used to non-parametrically recover unobserved productivity, Buettner (2005) has to restrict the transition probabilities $P(\omega_{jt}|\omega_{jt-1}, r_{jt-1})$ of the Markov process that governs the evolution of productivity to be of the form $P(\omega_{jt}|\psi_{jt})$, where $\psi_{jt} = \psi(\omega_{jt-1}, r_{jt-1})$ is an index that orders the probability distributions for ω_{jt}. The restriction to an index excludes the possibility that current productivity and R&D expenditures affect future productivity in qualitatively different ways.

To free up the relationship between current productivity, R&D expenditures, and future productivity, we base our empirical strategy on the firm's decisions on static (or "variable") inputs that are subsumed in the indirect profit function $\Pi(\cdot)$. Specifically, we follow LP and assume that labour l_{jt} and materials m_{jt} are chosen to maximize short-run profits.[6] This gives rise to two input

---

5. One can, *e.g.* combine the cost functions $C_i(\cdot)$ and $C_r(\cdot)$ into a more general cost function $C_{ir}(\cdot)$.

6. LP invoke this assumption to establish in their Equation (9) a sufficient condition for the invertibility of the intermediate input: on p. 320 just below Equation (1) LP assume that labour is "freely variable," on p. 322 just above

demand functions $l(s_{jt})$ and $m(s_{jt})$ that allow us to parametrically recover unobserved productivity and control for it in estimating the parameters of the production function in Equation (1) and the law of motion for productivity in Equation (3). Before detailing our empirical strategy in Section 4, we describe the data at hand.

3. DATA

Our data come from the Encuesta Sobre Estrategias Empresariales (ESEE) survey, a firm-level survey of Spanish manufacturing sponsored by the Ministry of Industry. The unit surveyed is the firm, not the plant or the establishment. At the beginning of this survey in 1990, 5% of firms with up to 200 workers were sampled randomly by industry and size strata. All firms with more than 200 workers were asked to participate, and 70% of these large firms chose to respond. Some firms vanish from the sample, due to either exit (shutdown by death or abandonment of activity) or attrition. The two reasons can be distinguished, and attrition remained within acceptable limits. To preserve representativeness, samples of newly created firms were added to the initial sample every year. Details on industry and variable definitions can be found in Appendix A.

Our sample covers a total of 1870 firms when restricted to firms with at least two years of data. Columns (1) and (2) of Table 1 show the number of observations and firms by industry. Samples sizes are moderate. Columns (3) and (4) show entry and exit. Newly created firms are a large share of the total number of firms, ranging from 15% to one-third in the different industries. In each industry there is a significant proportion of exiting firms (from 5% to above 10% in a few cases). Firms tend to remain in the sample for short periods, ranging from a minimum of two years to a maximum of 10 years between 1990 and 1999.

The 1990s were a period of rapid output growth, coupled with stagnant or at best slightly increasing employment and intense investment in physical capital, as can be seen from Columns (5)–(8) of Table 1. Consistent with this rapid growth, firms on average report that their markets are slightly more often expanding than contracting; hence, demand tends to shift out over time.

The growth of prices, averaged from the growth of prices as reported individually by each firm, is moderate. The growth of the price of output in Column (9) ranges from 0.7% to 2.2%. The growth of the wage and the price of materials ranges from 4.4% to 6.0% and, respectively, from 2.5% to 3.9%.

Our empirical strategy treats labour as a static input. This is appropriate because Spain greatly enhanced the possibilities for hiring and firing temporary workers during the 1980s and, by the beginning of the 1990s, had one the highest shares of temporary workers in Europe (Dolado *et al.*, 2002). In our sample the share of temporary workers ranges from 14% to 33% across industries. Rapid expansions and contractions of temporary workers are common: the difference between the maximum and the minimum share of temporary workers within a firm ranges on average from 12% to 20% across industries. In addition, we measure labour as hours worked (see Appendix A for details). At this margin at least firms enjoy a high degree of flexibility in determining the demand for labour: within a firm the difference between the maximum and the minimum hours worked ranges on average from 29% to 41%. These sizeable adjustments in hours worked are due to changes in hours worked per worker,[7] changes in the number of temporary workers, and changes in the number of permanent workers.

Equation (6) they assume that the intermediate input is also "freely variable," and they invoke short-run profit maximization at the start of the proof on p. 339.

7. The difference between the maximum and the minimum hours worked per worker within a firm ranges on average from 4% to 13% across industries.

TABLE 1
Descriptive statistics

Industry	Obs.[a] (1)	Firms[a] (2)	Entry[a] (%) (3)	Exit[a] (%) (4)	Rates of growth[a]					With R&D[b]			
					Output (s.d.) (5)	Capital (s.d.) (7)	Labour (s.d.) (6)	Materials (s.d.) (8)	Price (s.d.) (9)	Obs. (%) (10)	Stable (%) (11)	Occas. (%) (12)	R&D inten. (s.d.) (13)
1. Metals and metal products	1235	289	88 (30.4)	17 (5.9)	0.050 (0.238)	0.086 (0.278)	0.010 (0.183)	0.038 (0.346)	0.012 (0.055)	420 (34.0)	63 (21.8)	72 (24.9)	0.0126 (0.0144)
2. Non-metallic minerals	621	131	20 (15.3)	15 (11.5)	0.037 (0.208)	0.062 (0.238)	−0.001 (0.141)	0.039 (0.308)	0.010 (0.059)	186 (30.0)	16 (12.2)	41 (31.3)	0.0100 (0.0211)
3. Chemical products	1218	275	64 (23.3)	15 (5.5)	0.068 (0.196)	0.093 (0.238)	0.007 (0.146)	0.054 (0.254)	0.007 (0.061)	672 (55.2)	124 (45.1)	55 (20.0)	0.0268 (0.0353)
4. Agric. and ind. machinery	576	132	36 (27.3)	6 (4.5)	0.059 (0.275)	0.078 (0.247)	0.010 (0.170)	0.046 (0.371)	0.013 (0.032)	322 (55.9)	52 (39.4)	35 (26.5)	0.0219 (0.0275)
6. Transport equipment	637	148	39 (26.4)	10 (6.8)	0.087 (0.354)	0.114 (0.255)	0.011 (0.207)	0.087 (0.431)	0.007 (0.037)	361 (56.7)	62 (41.9)	35 (23.6)	0.0224 (0.0345)
7. Food, drink, and tobacco	1408	304	47 (15.5)	22 (7.2)	0.025 (0.224)	0.094 (0.271)	−0.003 (0.186)	0.019 (0.305)	0.022 (0.065)	386 (27.4)	56 (18.4)	64 (21.1)	0.0071 (0.0281)
8. Textile, leather, and shoes	1278	293	77 (26.3)	49 (16.7)	0.020 (0.233)	0.059 (0.235)	−0.007 (0.192)	0.012 (0.356)	0.016 (0.040)	378 (29.6)	39 (13.3)	66 (22.5)	0.0152 (0.0219)
9. Timber and furniture	569	138	52 (37.7)	18 (13.0)	0.038 (0.278)	0.077 (0.257)	0.014 (0.210)	0.029 (0.379)	0.020 (0.035)	66 (12.6)	7 (5.1)	18 (13.8)	0.0138 (0.0326)
10. Paper and printing products	665	160	42 (26.3)	10 (6.3)	0.035 (0.183)	0.099 (0.303)	−0.000 (0.140)	0.026 (0.265)	0.019 (0.089)	113 (17.0)	21 (13.1)	25 (13.8)	0.0143 (0.0250)

[a] Computed for 1991 to 1999 excluding the first observation for each firm.
[b] Computed for 1990 to 1998 excluding the last observation for each firm.

The R&D intensity of Spanish manufacturing firms is low by European standards, but R&D became increasingly important during the 1990s (see, *e.g.* European Commission, 2001).[8] The manufacturing sector consists partly of multinational companies with production facilities in Spain and very large R&D expenditures and partly of small and medium-sized companies that invested heavily in R&D in a struggle to increase their competitiveness in a growing and already very open economy.[9,10]

Columns (10)–(13) of Table 1 reveal that the nine industries differ markedly in terms of firms' R&D activities. Chemical products (3), agricultural and industrial machinery (4), and transport equipment (6) exhibit high innovative activity. The share of firms that perform R&D during at least one year in the sample period is about two-thirds, with slightly more than 40% of stable performers that engage in R&D in all years (Column (11)) and slightly more than 20% of occasional performers that engage in R&D in some (but not all) years (Column (12)). The average R&D intensity among performers ranges from 2.2% to 2.7% (Column (13)). The standard deviation of R&D intensity is substantial and shows that firms engage in R&D to various degrees and quite possibly with many different specific innovative activities. Metals and metal products (1), non-metallic minerals (2), food, drink and tobacco (7), and textile, leather, and shoes (8) are in an intermediate position. The share of firms that perform R&D is below one half and there are fewer stable than occasional performers. The average R&D intensity is between 1.0% and 1.5% with a much lower value of 0.7% in industry 7. Finally, timber and furniture (9) and paper and printing products (10) exhibit low innovative activity. The share of firms that perform R&D is around one quarter and the average R&D intensity is 1.4%.

4. EMPIRICAL STRATEGY

Our estimator builds on the insight of LP that the demands for static inputs such as labour and materials can be used to recover unobserved productivity. Static inputs are chosen with current productivity known and therefore contain information about it. Importantly, the demands for static inputs are the solution to the firm's short-run profit maximization problem. We are thus able to back out productivity without making assumptions on the firm's dynamic programming problem. Our estimator differs from LP by recognizing that, given a parametric specification of the production function, the functional form of the input demand functions (and their inverses) is known. Exploiting the parameter restrictions between the production function and the inverse input demand functions allows us to parameterically recover unobserved productivity.

To allow the firm to have some market power, *e.g.* because products are differentiated, we assume that the firm faces a downward sloping demand function that depends on the price of output p_{jt} and the demand shifter d_{jt}. Profit maximization requires that the firm sets the price that equates marginal cost to marginal revenue $P_{jt}\left(1 - \frac{1}{\eta(p_{jt},d_{jt})}\right)$, where $\eta(\cdot)$ is the absolute value

---

8. R&D intensities for manufacturing firms are 2.1% in France, 2.6% in Germany, and 2.2% in the UK as compared to 0.6% in Spain (European Commission, 2004).

9. At most a small fraction of the firms that engaged in R&D received subsidies that typically covered between 20% and 50% of R&D expenditures. The impact of subsidies is mostly limited to the amount that they add to the project, without crowding out private funds (see Gonzalez *et al.*, 2005). This suggests that R&D expenditures irrespective of their origin are the relevant variable for explaining productivity.

10. While some R&D expenditures were tax deductible during the 1990s, the schedule was not overly generous and most firms simply ignored it. A large reform that introduced some real stimulus took place towards the end of our sample period in 1999.

of the elasticity of demand.[11] Given the Cobb–Douglas production function in Equation (1), the assumption that labour and materials are static inputs implies that the labour demand function is

$$l(s_{jt}) = \frac{1}{1 - \beta_l - \beta_m} \Big(\beta_0 + (1 - \beta_m)\ln\beta_l + \beta_m\ln\beta_m + \mu + \beta_t t + \beta_k k_{jt} + \omega_{jt} - (1 - \beta_m)(w_{jt} - p_{jt})$$

$$- \beta_m(p_{Mjt} - p_{jt}) + \ln\Big(1 - \frac{1}{\eta(p_{jt}, d_{jt})}\Big)\Big), \tag{4}$$

where $\mu = \ln E\big[\exp(e_{jt})\big]$, $(w_{jt} - p_{jt})$ is the real wage, and $(p_{Mjt} - p_{jt})$ is the real price of materials. Solving Equation (4) for ω_{jt} we obtain the inverse labour demand function

$$h_l(t, k_{jt}, l_{jt}, p_{jt}, w_{jt}, p_{Mjt}, d_{jt}) = \lambda_l - \beta_t t - \beta_k k_{jt} + (1 - \beta_l - \beta_m)l_{jt} + (1 - \beta_m)(w_{jt} - p_{jt})$$

$$+ \beta_m(p_{Mjt} - p_{jt}) - \ln\Big(1 - \frac{1}{\eta(p_{jt}, d_{jt})}\Big), \tag{5}$$

where $\lambda_l = -\beta_0 - (1 - \beta_m)\ln\beta_l - \beta_m\ln\beta_m - \mu$ combines constants. These derivations generalize beyond the Cobb–Douglas production function in Equation (1): maintaining the Hicks-neutrality of productivity, the inverse labour demand function can be characterized in closed form for a CES or a generalized linear production function, although not for a translog production function.

To obtain our estimation equation, we proceed in two steps. We first substitute the law of motion for the controlled Markov process in Equation (3) into the production function in Equation (1). Then we use the inverse labour demand function $h_l(\cdot)$ in Equation (5) to substitute for ω_{jt-1} to obtain

$$y_{jt} = \beta_0 + \beta_t t + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + g(h_{ljt-1}, r_{jt-1}) + \xi_{jt} + e_{jt}, \tag{6}$$

where h_{ljt-1} is shorthand for $h_l(t-1, k_{jt-1}, l_{jt-1}, p_{jt-1}, w_{jt-1}, p_{Mjt-1}, d_{jt-1})$.[12] In Section 5 we also provide estimates using the inverse materials demand function $h_m(\cdot)$, although these have problems.

We base estimation on the moment restrictions

$$E\big[A(z_{jt})(\xi_{jt} + e_{jt})\big] = 0,$$

where $A(\cdot)$ is a vector of functions (to be specified in Section 4.2) of the exogenous variables z_{jt}.[13] Because ξ_{jt} is the innovation to productivity in period t, its value is not known to the firm

---

11. More generally the elasticity of the residual demand that a firm faces depends additionally on its rivals' prices. One may be able to replace rivals' prices by an aggregate price index or dummies, although this substantially increases the number of parameters to be estimated.

12. Our model nests a restricted version of the dynamic panel model of Blundell and Bond (2000). Leaving aside the fixed effects in their specification, we obtain their estimation Equation (2.2) by switching from a gross-output to a value-added production function with $\beta_m = 0$, restricting the law of motion in Equation (3) to an exogenous $AR(1)$ process with $g(\omega_{jt-1}) = \rho\omega_{jt-1}$, and using the marginal productivity condition for profit maximization to substitute y_{jt-1} for $-\ln\beta_l - \mu + l_{jt-1} + (w_{jt-1} - p_{jt-1}) - \ln\Big(1 - \frac{1}{\eta(p_{jt-1}, d_{jt-1})}\Big)$ in Equation (6). Hence, the differences between their and our approach lie in the generality of the productivity process and the strategy of estimation.

13. These moments in $\xi_{jt} + e_{jt}$ correspond to the "second-stage moments" in OP, LP, and ACF. The literature following OP also uses "first-stage moments" in e_{jt} that are obtained by using $h_l(\cdot)$ to substitute for ω_{jt} in the production function in Equation (1). Estimating this marginal productivity condition for profit maximization together with Equation (6) may increase efficiency. Parameters of interest may, however, cancel. In the special case of a value-added production function with $\beta_m = 0$ and perfect competition in the product market with $\eta(\cdot) = \infty$, the estimation equation is $y_{jt} = \lambda_1 + l_{jt} + (w_{jt} - p_{jt}) + e_{jt}$, where $\lambda_1 = -\ln\beta_l - \mu$ combines constants.

when it makes its decisions at time $t-1$. Hence, all lagged variables appearing in Equation (6) are uncorrelated with ξ_{jt}. Moreover, k_{jt}, whose value is determined in period $t-1$ by $i_{jt-1}=i(s_{jt-1})$, is uncorrelated with ξ_{jt} by virtue of our timing assumptions. Only l_{jt} and m_{jt} are correlated with ξ_{jt} (since ξ_{jt} is part of ω_{jt} and $l_{jt}=l(s_{jt})$ and $m_{jt}=m(s_{jt})$ are functions of ω_{jt}) and as endogenous variables must be instrumented for.

In considering instruments, it is important to remember that because Equation (6) models the law of motion for productivity it has an advantage over Equation (1): instruments have to be uncorrelated with the innovation to productivity ξ_{jt} but not necessarily with the level of productivity ω_{jt}. For example, while l_{jt-1} is uncorrelated with ξ_{jt} in Equation (6), l_{jt-1} is correlated with ω_{jt} in Equation (1) as long as productivity is correlated over time. Modelling the law of motion in this way mitigates the endogeneity problem in estimating production functions.

Our instrumenting strategy is grounded in a simultaneous-equations perspective. Equations (1), (3), (4), and (5) define a simultaneous-equations model. To simplify the exposition, we restrict the production function in Equation (1) to $y_{jt}=\beta_l l_{jt}+\omega_{jt}+e_{jt}$, the law of motion in Equation (3) to an exogenous Markov process $\omega_{jt}=g(\omega_{jt-1})+\xi_{jt}$, and assume perfect competition in the product market with $\eta(\cdot)=\infty$ to obtain the simultaneous-equations model

$$y_{jt}=\beta_l l_{jt}+g(\lambda_l+(1-\beta_l)l_{jt-1}+(w_{jt-1}-p_{jt-1}))+\xi_{jt}+e_{jt}, \tag{7}$$

$$l_{jt}=\frac{1}{1-\beta_l}\left(-\lambda_l+g(\lambda_l+(1-\beta_l)l_{jt-1}+(w_{jt-1}-p_{jt-1}))-(w_{jt}-p_{jt})+\xi_{jt}\right), \tag{8}$$

where $\lambda_l=-\ln\beta_l-\mu$ combines constants. The endogenous variables are y_{jt} and l_{jt}. While Equation (8) is in reduced form, Equation (7) is structural. To derive its reduced form, we substitute Equation (8) into Equation (7) to obtain

$$y_{jt}=\frac{1}{1-\beta_l}\left(-\beta_l\lambda_l+g(\lambda_l+(1-\beta_l)l_{jt-1}+(w_{jt-1}-p_{jt-1}))-\beta_l(w_{jt}-p_{jt})+\xi_{jt}\right)+e_{jt}. \tag{9}$$

The exogenous variables z_{jt} on the right-hand side of the reduced-form Equations (8) and (9) are the constant, $(w_{jt}-p_{jt})$, l_{jt-1}, and $(w_{jt-1}-p_{jt-1})$.

A potential concern is that the current real wage $(w_{jt}-p_{jt})$ and, by analogy, the current real price of materials $(p_{Mjt}-p_{jt})$ may be endogenous in the sense of being correlated with the innovation to productivity ξ_{jt}. Underlying OP, LP, and ACF is the reasoning that a change in productivity that is not anticipated by the firm is not correlated with its past decisions. This implies that lagged values are less susceptible to endogeneity than current values. In what follows we therefore restrict ourselves to the lagged real wage $(w_{jt-1}-p_{jt-1})$ and the lagged real price of materials $(p_{Mjt-1}-p_{jt-1})$ for instruments.

An important conclusion from viewing our model as a simultaneous-equations model is that the model is identified even in the simplest case of an exogenous $AR(1)$ process with $g(\omega_{jt-1})=\rho\omega_{jt-1}$ in Equation (7): we have three exogenous variables, namely the constant, l_{jt-1}, and $(w_{jt-1}-p_{jt-1})$, to estimate the three parameters $\lambda_0=\rho\lambda_l$, β_l, and ρ. Note that identification relies on the parameter restrictions between the production function and the inverse labour demand function.

Another conclusion is that Equation (7) does not suffice to estimate the parameters of the model absent price variation. Without variation in the real wage we are left with the constant

and l_{jt-1} as exogenous variables. Clearly, we cannot hope to estimate three parameters with two exogenous variables.[14]

Our empirical application is more elaborate than this example. First, we allow for imperfect competition, where the firm sets the current price of output p_{jt} in light of the level of productivity ω_{jt}. As ξ_{jt} is part of ω_{jt}, p_{jt} is endogenous. Imperfect competition is therefore an additional reason for not relying on the current real wage $(w_{jt} - p_{jt})$ and the current real price of materials $(p_{Mjt} - p_{jt})$ for instruments. In our empirical application we thus rely on the lagged price of output p_{jt-1} (as it separately enters our estimation Equation (6) through h_{ljt-1}), the lagged real wage $(w_{jt-1} - p_{jt-1})$, and the lagged real price of materials $(p_{Mjt-1} - p_{jt-1})$ for instruments; the tests for overidentifying restrictions in Section 5.1 indicate that the variation in p_{jt-1}, $(w_{jt-1} - p_{jt-1})$, and $(p_{Mjt-1} - p_{jt-1})$ is exogenous with respect to ξ_{jt} and therefore useful in estimating Equation (6).

Second, we allow the law of motion for productivity in Equation (3) to depend on lagged R&D expenditures r_{jt-1}. While r_{jt-1} is correlated with ω_{jt-1} (as $r_{jt-1} = r(s_{jt-1})$ by virtue of the policy function) and ω_{jt} (as productivity is correlated over time), our timing assumptions ensure that r_{jt-1} is uncorrelated with ξ_{jt} in Equation (6). Relying on r_{jt-1} for instruments allows us to estimate the impact of R&D on productivity. The policy function implies that r_{jt-1} varies with the lagged state variables s_{jt-1}, including the lagged nominal wage w_{jt-1} and the lagged nominal price of materials p_{Mjt-1}. Moreover, if a cost shifter x_{jt} is added to the cost functions $C_i(\cdot)$ and $C_r(\cdot)$ in Equation (2) to capture variation in investment opportunities, the price of equipment goods, and the nature of the undertaken project, then r_{jt-1} also varies with x_{jt-1}. All these variables are therefore a source of exogenous variation in R&D expenditures.

After more formally discussing identification in Section 4.1 we provide further details on our instrumenting strategy in Section 4.2. In Section 4.3, we then turn to the advantages and disadvantages of our parametric inversion.

4.1. *Identification*

Our estimation Equation (6) is a semiparametric, so-called partially linear, model with the additional restriction that the inverse labour demand function $h_l(\cdot)$ is of known form. In more formally discussing identification, we ask if we can separate the parametric from the non-parametric parts of the model. The importance of this question has been highlighted by ACF in their critique of OP and LP.

The fundamental condition for identification in a partially linear model is that the variables in the parametric part of the model are not perfectly predictable (in the least squares sense) by the variables in the non-parametric part (Robinson, 1988). In other words, there cannot be a functional relationship between the variables in the parametric and non-parametric parts (see Newey *et al.*, 1999).

To see how the known form of $h_l(\cdot)$ aids identification, suppose to the contrary that $h_l(\cdot)$ is of unknown form. In this case, the composition of $h_l(\cdot)$ and $g(\cdot)$ is another function of unknown form. The fundamental identification condition is violated if $h_l(\cdot)$ is of unknown form because k_{jt}

---

14. It may be possible to derive other estimation equations that allow us to estimate the parameters of the model absent price variation. Assume we observe revenue $(p + y_{jt})$ and the wage bill $(w + l_{jt})$ and consider the system of equations

$$(p + y_{jt}) = \lambda_1 + (w + l_{jt}) + e_{jt}, \tag{10}$$

$$(p + y_{jt}) = \lambda_2 + \beta_l(w + l_{jt}) + \rho(1 - \beta_l)(w + l_{jt-1}) + \xi_{jt} + e_{jt}, \tag{11}$$

where $\lambda_1 = -\ln\beta_l - \mu$ and $\lambda_2 = -\rho(\ln\beta_l + \mu) - (1 - \rho)(\beta_l w - p)$ combine constants. In a Monte Carlo study we first estimated Equation (10) by OLS to obtain λ_1 and then recovered β_l from λ_1 using the definition $\mu = \ln E[\exp(e_{jt})]$. Finally, taking β_l as given, we estimated Equation (11) by OLS to obtain λ_2 and ρ.

is perfectly predictable from the variables in the non-parametric part. To see this, recall that $K_{jt} = (1-\delta)K_{jt-1} + \exp(i(s_{jt-1}))$ by the law of motion and the policy function for investment in physical capital. The central question is whether k_{jt-1} and $s_{jt-1} = (t-1, k_{jt-1}, \omega_{jt-1}, w_{jt-1}, p_{Mjt-1}, d_{jt-1})$ can be inferred from $(t-1, k_{jt-1}, l_{jt-1}, p_{jt-1}, w_{jt-1}, p_{Mjt-1}, d_{jt-1}, r_{jt-1})$, the arguments of the composition of $h_l(\cdot)$ and $g(\cdot)$. The answer is affirmative because ω_{jt-1} is by construction a function of all arguments of $h_l(\cdot)$. Hence, there is a collinearity problem similar to the one that ACF ascertain for the estimator in LP.[15]

Our approach differs from LP in that it exploits the known form of the inverse labour demand function $h_l(\cdot)$. In this case, the central question becomes whether k_{jt-1} and s_{jt-1} can be inferred from h_{ljt-1}, the value of $h_l(\cdot)$ as opposed to its arguments, and r_{jt-1}. This is not the case: while h_{ljt-1} is identical to ω_{jt-1}, we cannot infer the remaining state variables from h_{ljt-1} and r_{jt-1}. In particular, even if $r_{jt-1} = r(s_{jt-1})$ happened to be invertible for, say, k_{jt-1} we are still short of w_{jt-1}, p_{Mjt-1}, and d_{jt-1}.[16] Since k_{jt} in the parametric part is not perfectly predictable from h_{ljt-1} and r_{jt-1} in the non-parametric part, the model is identified. Note that this argument rests on either the wage, the price of materials, or the demand shifter being a state variable in the firm's dynamic programming problem.

4.2. *Estimation*

Define the residual of our estimation Equation (6) as a function of the parameters θ to be estimated as

$$v_{jt}(\theta) = y_{jt} - \left(\beta_0 + \beta_t t + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + g(h_{ljt-1}, r_{jt-1})\right).$$

The GMM problem is

$$\min_{\theta} \left[\frac{1}{N}\sum_j A(z_j)v_j(\theta)\right]' W_N \left[\frac{1}{N}\sum_j A(z_j)v_j(\theta)\right], \tag{12}$$

where $A(\cdot)$ is a $L \times T_j$ matrix of functions of the exogenous variables z_j and $v_j(\cdot)$ is a $T_j \times 1$ vector with L being the number of instruments, T_j the number of observations of firm j, and N the number of firms. We use the two-step GMM estimator of Hansen (1982). We first obtain a consistent estimate $\hat{\theta}$ of θ with weighting matrix $W_N = \left(\frac{1}{N}\sum_j A(z_j)A(z_j)'\right)^{-1}$. This first step is the NL2SLS estimator of Amemiya (1974). In the second step we then compute the optimal estimate with weighting matrix $W_N = \left(\frac{1}{N}\sum_j A(z_j)v_j(\hat{\theta})v_j(\hat{\theta})'A(z_j)'\right)^{-1}$.

Markov process: We allow for a different function when the firm adopts the corner solution of zero R&D expenditures and when it chooses positive R&D expenditures and specify $g(h_{ljt-1}, r_{jt-1})$ as

$$1(R_{jt-1}=0)\left(g_{00} + g_{01}(h_{ljt-1} - \lambda_l)\right) + 1(R_{jt-1}>0)\left(g_{10} + g_{11}(h_{ljt-1} - \lambda_l, r_{jt-1})\right). \tag{13}$$

---

15. The fundamental identification condition is, however, satisfied if a cost shifter x_{jt} is added to $C_i(\cdot)$ and $C_r(\cdot)$: while $i(\cdot)$ depends on x_{jt}, x_{jt} cannot be inferred from the arguments of the composition of $h_l(\cdot)$ and $g(\cdot)$. Alternatively, one may follow OP, LP, and ACF and invoke "first-stage moments" in e_{jt}.

16. In our application we have a substantial fraction of observations with zero R&D expenditures (between 43% and 83% depending on the industry, see Column (10) of Table 1) that cannot possibly be inverted for the stock of capital. More generally, we suspect that R&D projects may often be lumpy and involve substantial and possibly non-convex adjustment costs that negate invertibility.

Since the constants g_{00} and g_{10} cannot be separately estimated from the constant β_0 in Equation (6), we estimate the combined constant $\lambda_0 = \beta_0 + g_{00}$ and include the dummy for performers $1(R_{jt-1} > 0)$ to measure $g_{10} - g_{00}$.

Series approximation: We model an unknown function $q(\cdot)$ of one variable u by a univariate polynomial of degree three. We model an unknown function $q(\cdot)$ of two variables u and v by a complete set of polynomials of degree three. We specify the absolute value of the elasticity of demand as $1 + \exp(q(p_{jt-1}, d_{jt-1}))$ in order to impose the theoretical restriction $\eta(p_{jt-1}, d_{jt-1}) > 1$.

Secular trend: After experimentation we omit the trend t in industries 4 and 9, where we tend to have less data. We model it by a linear time trend in industries 2, 3, 6, 8, and 10 and by dummies in industries 1 and 7, where we tend to have more data.

Parameters: Our baseline specification with time trend has 27 parameters: constant, time trend, three production function coefficients, thirteen coefficients in the series approximation of $g(\cdot)$, and nine coefficients in the series approximation of $\eta(\cdot)$.

Instruments: The literature on optimal instruments shows that setting

$$A(z_j) = E\left[\left.\frac{\partial v_j(\theta_0)'}{\partial \theta}\right| z_j\right],\tag{14}$$

where θ_0 is the true value of θ, minimizes the variance of the GMM estimator (Amemiya, 1974; Newey, 1990, 1993). For general non-linear models the conditional expectation in Equation (14) is difficult or perhaps even impossible to compute. Substituting a non-parametric estimate of it into the GMM problem is feasible but cumbersome.[17]

Fortunately, however, our model affords a more direct approach because our estimation Equation (6) is linear in the endogenous variables and the reduced-form equations are available in closed form. Returning for simplicity to the example in Equations (7)–(9), the conditional expectation of the derivative of $v_{jt}(\theta)$ with respect to β_l, say, is

$$-\frac{1}{1-\beta_l}\left(-\lambda_l + g(\lambda_l + (1-\beta_l)l_{jt-1} + (w_{jt-1} - p_{jt-1})) - (w_{jt} - p_{jt})\right)$$

$$+ l_{jt-1}\frac{\partial g(\lambda_l + (1-\beta_l)l_{jt-1} + (w_{jt-1} - p_{jt-1}))}{\partial \omega_{jt-1}}.$$

Because we approximate the conditional expectation function $g(\cdot)$ by a univariate polynomial, this expression is a linear combination of the constant, $(w_{jt} - p_{jt})$, and the complete set of polynomials in l_{jt-1} and $(w_{jt-1} - p_{jt-1})$. Using the constant, $(w_{jt} - p_{jt})$, and the complete set of polynomials in l_{jt-1} and $(w_{jt-1} - p_{jt-1})$ as instruments therefore achieves the same variance as using optimal instruments (Arellano, 2003, p. 206).

In our empirical application we similarly use polynomials in the exogenous variables as instruments. This strategy is widely employed in the literature, *e.g.* in Wooldridge's (2009) GMM version of the two-stage procedures of OP, LP, and ACF and in the sieve estimation procedure of Ai and Chen (2003, 2007); see also the discussion on p. 862 of Berry *et al.* (1995). The exogenous variables we rely on are the constant, t, k_{jt}, k_{jt-1}, l_{jt-1}, m_{jt-1}, p_{jt-1}, $(w_{jt-1} - p_{jt-1})$, $(p_{Mjt-1} -$

---

17. Equation (14) presumes $\Omega(z_j) = E\left[v_j(\theta_0)v_j(\theta_0)' | z_j\right] = \sigma^2 I_{T_j}$. Instead of non-parametrically estimating $\Omega(z_j)$ along with $A(z_j)$ we use the two-step GMM estimator of Hansen (1982) to account for heteroskedasticity and autocorrelation of unknown form.

p_{jt-1}), d_{jt-1}, r_{jt-1}, and $1(R_{jt-1}>0)=1-1(R_{jt-1}=0)$. We omit p_{jt} because we consider an imperfectly competitive setting and d_{jt} because it is highly correlated with d_{jt-1}. Finally, we omit $(w_{jt}-p_{jt})$ and $(p_{Mjt}-p_{jt})$ to guard against potential endogeneity concerns.

In light of Equation (13) spanning the optimal instruments requires a complete set of polynomials in the arguments of $h_l(\cdot)$ interacted with $1(R_{jt-1}=0)$ and a complete set of polynomials in the arguments of $h_l(\cdot)$ and r_{jt-1} interacted with $1(R_{jt-1}>0)$. To arrive at a number of overidentifying restrictions that is reasonable for the data at hand, we select after experimentation the most important terms for predicting the derivative in Equation (14).

For industries 4, 9, and 10 we use a basic set of 69 instruments: the constant, t, k_{jt}, m_{jt-1}, $1(R_{jt-1}>0)$ (5 instruments); a complete set of polynomials in p_{jt-1} and d_{jt-1} (9 instruments); a complete set of polynomials in k_{jt-1}, l_{jt-1}, $(w_{jt-1}-p_{jt-1})$, and $(p_{Mjt-1}-p_{jt-1})$ (34 instruments); a univariate polynomial in r_{jt-1} (3 instruments); and all interactions of degree less than three of r_{jt-1} with k_{jt-1}, l_{jt-1}, $(w_{jt-1}-p_{jt-1})$, $(p_{Mjt-1}-p_{jt-1})$, p_{jt-1}, and d_{jt-1} (18 instruments).

In the remaining industries, where we tend to have more data, we use additional instruments: in industries 2, 3, and 6 we add the interactions of $1(R_{jt-1}>0)$ with k_{jt-1}, l_{jt-1}, m_{jt-1}, $(w_{jt-1}-p_{jt-1})$, and $(p_{Mjt-1}-p_{jt-1})$ (5 instruments); in industry 7 we add the interactions of $t-1$ with k_{jt-1}, l_{jt-1}, $(w_{jt-1}-p_{jt-1})$, and $(p_{Mjt-1}-p_{jt-1})$ (4 instruments); and in industries 1 and 8 we do both.

Implementation: Gauss code for our estimator along with instructions for obtaining the data are available as supplementary data. To reduce the complexity of the GMM problem in Equation (12), we "concentrate out" the constant and the thirteen coefficients in the series approximation of $g(\cdot)$ that enter it linearly (Wooldridge, 2010, p. 435). To guard against local minima we have extensively searched over the remaining parameters, often using preliminary estimates to narrow down the range of these parameters.

Productivity estimates: Once the model is estimated, we can recover actual productivity h_{ljt} and expected productivity $g(\cdot)$ up to a constant. We estimate the productivity innovation ξ_{jt} up to a constant as the difference between h_{ljt} and $g(\cdot)$. Taken together, $g(\cdot)$ and ξ_{jt} characterize the controlled Markov process for productivity. We can also estimate the random shocks e_{jt}. In the remainder of this article we let $\widehat{h}_{ljt}$, $\widehat{g}(\cdot)$, $\widehat{\xi}_{jt}$, and $\widehat{e}_{jt}$ denote these estimates.

4.3. *Parametric vs. non-parametric inversion*

We differ from the literature following OP by recognizing that the parametric specification of the production function in combination with the assumption that labour and materials are static inputs implies a known form for the inverse labour demand function $h_l(\cdot)$. We can thus parametrically recover unobserved productivity and control for it in estimating the parameters of the production function in Equation (1) and the law of motion for productivity in Equation (3).

A seeming drawback of our parametric approach is that it requires firm-level wage and price data. However, the same is true for a non-parametric approach: the demand for labour is a function of the wage and prices whether one inverts it parametrically or non-parametrically (see the discussion on p. 323 of LP); by spelling out the demand for labour in Equation (4) our parametric approach just makes the role of the wage and prices explicit.

In the absence of firm-level wage and price data, one may be able to replace the wage and prices in the firm's short-run profit maximization problem by dummies (as in LP and ACF) or aggregate wage and price indices. This may be justified if firms can be assumed to operate in identical environments because inputs and output are homogenous or symmetrically differentiated. One may then also have to confront an issue raised by Bond and Söderbom (2005). They argue that,

absent any variation in the wage and prices, it may be hard to estimate the coefficients on static inputs in a Cobb–Douglas production function.

Making full use of the structural assumptions by parametrically recovering unobserved productivity aids identification (Section 4.1). Moreover, we have but a single equation to estimate, and only the conditional expectation function $g(\cdot)$ is unknown and must be estimated non-parametrically (Section 4.2). This yields a particularly simple estimator for production functions. In contrast, the previous literature either relies on a two-stage procedure (OP, LP, and ACF) or jointly estimates a system of equations (Wooldridge, 2009). The fact that only the conditional expectation function $g(\cdot)$ is unknown and must be estimated non-parametrically means that our estimator can be applied even when the available data is insufficient to non-parametrically estimate an inverse labour demand function $h_l(\cdot)$ with more than just a few arguments.

Our parametric approach rests on the assumption that the demands for static inputs are the solution to the firm's short-run profit maximization problem. This assumption may or may not be valid in a given application. However, as detailed in ACF, the non-parametric approach also relies on specific assumptions that may or may not be valid in a given application.

The non-parametric approach in OP, LP, and ACF extends to our model of endogenous productivity change.[18] The advantage is that it can accommodate situations where the assumption of short-run profit maximization is violated. There are at least two difficulties, however. First, if input choices have dynamic implications say because of adjustment costs, then lagged input choices are typically state variables that must be accounted for in non-parametrically recovering unobserved productivity. Second, to the extent that the non-parametric approach relies on inverting decisions that have dynamic implications (as in OP and parts of ACF), it requires the researcher to first prove invertibility from the firm's dynamic programming problem, and the discussion of Buettner (2005) in Section 2 suggests that this can be difficult especially if the productivity process is endogenized. We thus view our parametric approach as complementing the existing literature for situations where one can be reasonably confident in the assumption of short-run profit-maximization.

5. PRODUCTION FUNCTION ESTIMATES AND COMPARISON TO KNOWLEDGE CAPITAL MODEL

We first present our estimates of the production function and the Markov process that governs the evolution of productivity. We next show that the link between R&D and productivity is subject to a high degree of non-linearity and uncertainty. Then we provide a more detailed comparison of our model of endogenous productivity change and the knowledge capital model.

5.1. *Production function estimates*

Table 2 summarizes different production function estimates. Columns (1)–(3) report OLS estimates for the coefficients of the Cobb-Douglas production function in Equation (1). The coefficients are reasonable and returns to scale, as given by $\beta_k + \beta_l + \beta_m$, are close to constant.

Columns (4)–(6) of Table 2 report GMM estimates for our leading specification in Equation (6). Compared to the OLS estimates, the changes go in the direction expected from theory and match the results in OP and LP. The labour coefficients decrease considerably in eight industries while the capital coefficients increase somewhat in six industries. The materials coefficients show no particular pattern.

---

18. Essentially all one has to do is replace $\omega_{jt} = g(\omega_{jt-1}) + \xi_{jt}$ by $\omega_{jt} = g(\omega_{jt-1}, r_{jt-1}) + \xi_{jt}$; everything else remains the same (see De Loecker, 2010).

TABLE 2
Production function estimates and specification tests

Industry	OLS[a] Capital (std. err.) (1)	Labour (std. err.) (2)	Materials (std. err.) (3)	GMM[a] Capital (std. err.) (4)	Labour (std. err.) (5)	Materials (std. err.) (6)	Overidentifying restrictions test $\chi^2(df)$ (7)	p val. (8)	Parameter restrictions test $\chi^2(3)$ (9)	p val. (10)
1. Metals and metal products	0.109 (0.013)	0.252 (0.022)	0.642 (0.020)	0.106 (0.014)	0.111 (0.031)	0.684 (0.011)	62.553 (51)	0.129	11.666	0.009
2. Non-metallic minerals	0.096 (0.021)	0.275 (0.034)	0.655 (0.028)	0.227 (0.014)	0.137 (0.016)	0.633 (0.014)	50.730 (47)	0.329	6.047	0.109
3. Chemical products	0.060 (0.010)	0.239 (0.021)	0.730 (0.020)	0.132 (0.015)	0.122 (0.026)	0.713 (0.011)	48.754 (47)	0.402	0.105	0.991
4. Agric. and ind. machinery	0.051 (0.017)	0.284 (0.038)	0.671 (0.027)	0.079 (0.015)	0.281 (0.029)	0.642 (0.013)	45.833 (44)	0.396	1.798	0.615
6. Transport equipment	0.080 (0.023)	0.289 (0.033)	0.636 (0.046)	0.117 (0.015)	0.158 (0.023)	0.675 (0.016)	40.296 (47)	0.745	0.414	0.937
7. Food, drink, and tobacco	0.094 (0.014)	0.177 (0.016)	0.739 (0.016)	0.068 (0.014)	0.129 (0.024)	0.766 (0.008)	61.070 (46)	0.068	8.866	0.031
8. Textile, leather, and shoes	0.059 (0.010)	0.335 (0.024)	0.605 (0.019)	0.057 (0.011)	0.313 (0.016)	0.593 (0.013)	66.143 (51)	0.075	4.749	0.191
9. Timber and furniture	0.079 (0.019)	0.283 (0.029)	0.670 (0.029)	0.131 (0.009)	0.176 (0.017)	0.697 (0.011)	44.951 (43)	0.390	0.618	0.892
10. Paper and printing products	0.092 (0.016)	0.321 (0.029)	0.621 (0.025)	0.121 (0.013)	0.249 (0.025)	0.617 (0.014)	51.371 (42)	0.152	5.920	0.118

[a] All standard errors are robust to heteroskedasticity and autocorrelation.

To assess the validity of our estimates we subject them to a battery of specification tests and robustness checks.[19]

Overidentifying restrictions: We first test for overidentifying restrictions or validity of the moment conditions.[20] At a 5% level the validity of the moment conditions cannot be rejected in any industry, see Columns (7) and (8) of Table 2.

The lagged wage and the lagged price of materials play a key role in the estimation. To more explicitly validate them as instruments we compute the difference in the value of the objective function for our leading specification to its value when the subset of moments involving either the lagged wage or the lagged price of materials are excluded.[21] The exogeneity assumption on the lagged wage is rejected in industry 8 at a 5% level while that on the lagged price of materials cannot be rejected in any industry.[22]

We similarly assess the validity of lagged labour, lagged materials, and current and lagged capital as instruments. While the exogeneity assumption on lagged labour cannot be rejected in any industry at a 5% level, that on lagged materials is rejected in industry 8. Finally, the exogeneity assumption on current and lagged capital is rejected at a 5% level in industries 2 and 10. Viewing all these tests in conjunction, however, we feel that there is little ground for concern regarding the validity of the moment conditions that we are using.

Imperfect competition: We test for perfect competition in the product market with $\eta(\cdot) = \infty$ by removing the function in the price of output p_{jt} and the demand shifter d_{jt} from $h_l(\cdot)$ in Equation (5).[23] The data reject perfect competition at a 5% level in seven industries and in all industries at a 10% level. Our estimates of the average elasticity of demand are around 2.

Parameter restrictions: The coefficients β_k, β_l, and β_m appear both in the production function in Equation (1) and in the inverse labour demand function in Equation (5). We test the implied restrictions on the parameters of our estimation Equation (6). As Columns (9) and (10) of Table 2 show, we reject at a 5% level that the coefficients in the production function equal their counterparts in the inverse labour demand function in industries 1 and 7.

---

19. We follow LP and ACF and do not account for sample selection by modelling a firm's exit decision because the institutional realities in Spain render it unlikely that a firm is able to exit the industry immediately after receiving an adverse shock to productivity (see, *e.g.* Djankov *et al.*, 2002). Moreover, regressing a dummy for exiting in the subsequent period on a dummy for engaging in R&D in the current period, time dummies, product submarket dummies, an index of market dynamism, a dummy for large firms with more than 200 workers, and a dummy for firms older than five years shows that being engaged in R&D is associated with a reduced probability of exit in many industries. This may be because innovative activities often imply large sunk cost that make a firm more reluctant to exit the industry or at least to exit it immediately, thus further mitigating the potential for sample selection.

20. The value of the GMM objective function for the optimal estimator, scaled by N, has a limiting χ^2 distribution with $L - P$ degrees of freedom, where L is the number of instruments and P the number of parameters to be estimated.

21. To use the same weighting matrix for both specifications, we delete the rows and columns corresponding to the excluded moments from the weighting matrix of the optimal estimator.

22. As a further check we have replaced the lagged firm-specific wage as an instrument by the lagged average earnings per hour of work in the manufacturing sector from the Encuesta de Salarios, an employee survey conducted by the Instituto Nacional de Estadistica. Compared to Columns (4)–(6) of Table 2, the standard errors tend to increase as expected from an instrument that does not vary across firms. The most visible changes are in the labour coefficient, which increases in industries 1, 3, 9, and 10 and decreases in industries 4, 6, and 8. The capital coefficient decreases in industries 3 and 9 and increases in industries 4, 6, 7, and 8. The materials coefficient remains essentially the same in all industries. The absence of systematic changes confirms that the variation in w_{jt-1} is exogenous with respect to ξ_{jt} and therefore useful in estimating Equation (6).

23. We test whether the model satisfies one or more restrictions by using the weighting matrix for the optimal estimator to compute the restricted estimator. The difference of the GMM objective functions, scaled by N, has a limiting χ^2 distribution with degrees of freedom equal to the number of restrictions. Similarly, we test whether one or more restrictions can be relaxed by using the weighting matrix for the optimal estimator to compute the unrestricted estimator.

There are many possible reasons. First, the functional form of the production function in Equation (1) may be inappropriate. However, the Cobb–Douglas production function is a first-order differential approximation to a general production function (Chambers, 1988, Sections 5.1 and 5.5).[24] Second, the functional form of the inverse labour demand function in Equation (5) may be inappropriate if the labour decision has dynamic consequences. In measuring labour we combine the hours worked by permanent workers with the hours worked by temporary workers. A concern may be that these types of workers differ in their productivity and the degree of flexibility that firms have in adjusting them. We return to this concern below. Third, the literature following OP rests on the assumption that unobserved productivity can be recovered from observed decisions without error. To this end, it rules out measurement error by way of the so-called scalar unobservable assumption (Ackerberg *et al.*, 2007, Section 2.3).[25] Our estimator is similarly vulnerable to measurement error. Indeed, as we show in Appendix B, imperfectly observable wages and prices may drive a wedge between the coefficients in the production function and their counterparts in the inverse labour demand function. In our view, measurement error is a plausible explanation for rejecting the parameter restrictions, especially because the existing literature has emphasized the importance of measurement error in practice (see, *e.g.* p. 326 of LP).

The parameter restrictions test is just one of many specification tests that we conduct. While in industries 1 and 7 there is a statistical tension between β_k, β_l, and β_m in the production function and in the inverse labour demand function, there is evidence that the scope of the problem is limited. First, if in industries 1 and 7 we test one coefficient at a time, then we reject at a 5% level that that coefficient in the production function equals its counterpart in the inverse labour demand function in one of six cases. This suggests that the gap for any individual coefficient tends to be small. Second, we no longer reject the parameter restrictions in industries 1 and 7 at a 5% level if we simply use more instruments. This suggests that the instruments we select for our leading specification do not fully account for all the non-linearities in the estimation Equation (6) so that a pattern remains in the residuals. Third, our conclusions regarding the link between productivity and R&D remain unchanged if we drop the parameter restrictions to consistently estimate the parameters of the production function and recover productivity in case wages and prices are imperfectly observable (see Appendix B for details).

Alternative inversion: Our model allows us to alternatively use labour or materials to recover productivity. The inverse materials demand function is

$$h_m(t,k_{jt},m_{jt},p_{jt},w_{jt},p_{Mjt},d_{jt})=\lambda_m-\beta_t t-\beta_k k_{jt}+(1-\beta_l-\beta_m)m_{jt}+\beta_l(w_{jt}-p_{jt})$$

$$+(1-\beta_l)(p_{Mjt}-p_{jt})-\ln\left(1-\frac{1}{\eta(p_{jt},d_{jt})}\right),$$

where $\lambda_m=-\beta_0-(1-\beta_l)\ln\beta_m-\beta_l\ln\beta_l-\mu$ combines constants.

Columns (1)–(3) of Table 3 report GMM estimates of the production function coefficients and Columns (4)–(6) specification tests when we use the inverse materials demand function $h_m(\cdot)$ instead of the inverse labour demand function $h_l(\cdot)$. The estimates using materials have problems. We reject the validity of the moment conditions in four industries at a 5% level. We further reject

---

24. This implies that the marginal productivity underlying the demand for labour in Equation (4) is a zero-order approximation. The usefulness of these approximations is limited by the fact that they hold for an arbitrary but fixed vector of inputs whereas in practice firms differ widely in the inputs they use.

25. The literature has just begun to relax this assumption by combining different inversions while allowing each of them to be subject to error (Huang and Hu, 2011).

TABLE 3

Production function estimates and specification tests: inverse materials demand function

Industry	GMM[a] Capital (std. err.)	Labour (std. err.)	Materials (std. err.)	Overidentifying restrictions test $\chi^2(df)$	p val.	Param. re. test $\chi^2(3)$ p val.	Non-nest. test $N(0,1)$ p val.	Equality of capital $\chi^2(1)$ p val.	Equality of sum of lab. and mat. $\chi^2(1)$ p val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Metals and metal products	0.123	0.247	0.543	70.63	0.036	10.604	−0.959	0.179	0.167
	(0.013)	(0.014)	(0.022)	(51)		0.014	0.169	0.672	0.683
2. Non-metallic minerals	0.123	0.315	0.604	51.73	0.294	0.855	0.592	23.540	36.398
	(0.014)	(0.012)	(0.018)	(47)		0.836	0.723	0.000	0.000
3. Chemical products	0.106	0.240	0.622	52.48	0.270	6.178	−0.688	0.741	1.525
	(0.017)	(0.012)	(0.024)	(47)		0.103	0.246	0.389	0.217
4. Agric. and ind. machinery	0.078	0.352	0.550	39.85	0.650	6.232	−0.389	0.400	1.028
	(0.013)	(0.014)	(0.019)	(44)		0.101	0.349	0.527	0.311
6. Transport equipment	0.139	0.239	0.553	59.59	0.103	10.415	0.000	1.264	2.448
	(0.016)	(0.016)	(0.022)	(47)		0.015	0.500	0.261	0.118
7. Food, drink, and tobacco	0.147	0.211	0.608	53.58	0.206	13.825	−2.169	20.024	8.941
	(0.017)	(0.011)	(0.025)	(46)		0.003	0.015	0.000	0.003
8. Textile, leather, and shoes	0.204	0.294	0.334	69.48	0.044	3.071	−0.701	32.347	36.605
	(0.025)	(0.012)	(0.041)	(51)		0.381	0.242	0.000	0.000
9. Timber and furniture	0.037	0.266	0.666	68.25	0.008	33.408	−1.298	26.224	60.464
	(0.015)	(0.013)	(0.018)	(43)		0.000	0.097	0.000	0.000
10. Paper and printing products	0.162	0.329	0.499	61.43	0.027	3.371	−2.109	3.496	1.124
	(0.013)	(0.016)	(0.016)	(42)		0.338	0.017	0.062	0.289

[a] All standard errors are robust to heteroskedasticity and autocorrelation.

the parameter restrictions in four industries. Finally, the capital coefficient in industry 9 and the materials coefficient in industry 8 are implausibly small.

To compare the overall goodness of fit we apply the Rivers and Vuong (2002) test for model selection among non-nested models.[26] As can be seen from Column (7) of Table 3, at a 5% level the data do not favour one inversion over another in seven industries. In the remaining two industries labour is favoured over materials.

Following LP we test whether the capital coefficient is the same when we use materials instead of labour to recover productivity. We cannot reject the equality of the capital coefficient in five industries at a 5% level, see Column (8) of Table 3. Moreover, the estimates using materials are especially problematic in three of the four industries where we reject.

To the extent that different inversions perform differently there is a gap between the model and the data-generating process. While pinpointing the exact source of this gap is difficult, our findings are consistent with materials being more prone to measurement error than labour due to inventories and, in particular, subcontracting and outsourcing. Over 40% of firms contract for parts and pieces with outside providers and, among these firms, outsourcing on average amounts to over 20% of the value of materials. A firm outsources to gain access to either cheaper or better parts and pieces than it can produce inhouse (Antràs and Helpman, 2004; Helpman *et al.*, 2004; Grossman and Rossi-Hansberg, 2008). Treating materials as a single input into the production process therefore introduces an additional unobservable in the form of the prices and productivities of subcontracted parts and pieces and those produced inhouse into the first-order conditions for static inputs. This additional unobservable may act like measurement error and cause us to reject the parameter restrictions test by driving a wedge between the coefficients in the production function and their counterparts in the materials labour demand function (see again Appendix B).

Comparing the labour and materials coefficients finally reveals an interesting pattern that the literature following OP may have overlooked. The labour coefficient tends to be higher and the materials coefficient lower when we use materials instead of labour to recover productivity. What appears to be reasonably well estimated is the sum of the labour and materials coefficients: we cannot reject the equality of the sum in five industries at a 5% level, see Column (9) of Table 3. This hints at collinearity. Its likely origin are the powers of the lagged input (l_{jt-1} or m_{jt-1}) that appear in the non-parametric part ($g(h_{ljt-1}, r_{jt-1})$ or $g(h_{Mjt-1}, r_{jt-1})$) of the estimation equation. It may be difficult to separate the impact of the lagged input on the non-parametric part from its impact on the production function.

Permanent and temporary workers: In measuring labour we combine the hours worked by permanent workers with the hours worked by temporary workers. The productivity of permanent workers may, however, differ from that of temporary workers and firms may incur a substantial

---

26. Consider models 1 and 2 and define $\hat{v}_{nj} = v_j(\hat{\theta}_n)$ and $A_{nj} = A_n(z_j)$, where $\hat{\theta}_n$ are the estimates for model n, $A_n(\cdot)$ is a matrix of functions of the exogenous variables z_j, and W_{nN} is the first-step weighting matrix. Then the difference between the GMM objective functions, scaled by $\sqrt{N}$, has an asymptotic normal distribution with zero mean and variance

$$\sigma^2 = 4\Bigg[\left(\frac{1}{N} \sum_j A_{1j} \hat{v}_{1j} \right)' W_{1N} \left(\frac{1}{N} \sum_j A_{1j} \hat{v}_{1j} \hat{v}'_{1j} A'_{1j} \right) W_{1N} \left(\frac{1}{N} \sum_j A_{1j} \hat{v}_{1j} \right)$$

$$+ \left(\frac{1}{N} \sum_j A_{2j} \hat{v}_{2j} \right)' W_{2N} \left(\frac{1}{N} \sum_j A_{2j} \hat{v}_{2j} \hat{v}'_{2j} A'_{2j} \right) W_{2N} \left(\frac{1}{N} \sum_j A_{2j} \hat{v}_{2j} \right)$$

$$-2 \left(\frac{1}{N} \sum_j A_{1j} \hat{v}_{1j} \right)' W_{1N} \left(\frac{1}{N} \sum_j A_{1j} \hat{v}_{1j} \hat{v}'_{2j} A'_{2j} \right) W_{2N} \left(\frac{1}{N} \sum_j A_{2j} \hat{v}_{2j} \right) \Bigg].$$

cost for hiring and firing permanent workers. In Appendix C we extend our model from Section 2 in two ways. First, we replace the production function in Equation (1) by

$$y_{jt} = \beta_0 + \beta_t t + \beta_k k_{jt} + \beta_l \alpha l_{Pjt} + \beta_l (1-\alpha) l_{Tjt} + \beta_m m_{jt} + \omega_{jt} + e_{jt}, \tag{15}$$

where l_{Pjt} and l_{Tjt} are the log of permanent and temporary labour, respectively, and α parameterizes productivity differences between permanent and temporary workers.[27] Second, we introduce adjustment costs $C_{L_P}(L_{Pjt}, L_{Pjt-1})$ into the Bellman Equation (2). To be able to retain our approach to estimation, we restrict ourselves to convex adjustment costs (Cooper and Willis, 2004).[28] As we show in Appendix C, the first-order conditions for permanent and temporary labour can be combined into an aggregate labour demand function

$$l(s_{jt}) + c_{1jt} = \frac{1}{1 - \beta_l - \beta_m} \Bigg(\beta_0 + (1-\beta_m)\ln\beta_l + \beta_m\ln\beta_m + \mu + \beta_t t + \beta_k k_{jt} + \omega_{jt}$$

$$- (1-\beta_m)(w_{jt} + c_{2jt} - c_{1jt} - p_{jt}) - \beta_m(p_{Mjt} - p_{jt}) + \ln\left(1 - \frac{1}{\eta(p_{jt}, d_{jt})}\right) \Bigg), \tag{16}$$

where

$$c_{1jt} = \alpha\ln(1 - S_{Tjt}) + (1-\alpha)\ln S_{Tjt}, \tag{17}$$

$$c_{2jt} = \ln\left(1 + \frac{(1 - S_{Tjt})\Delta_{jt}}{1 + \left(\frac{W_{Tjt} - W_{Pjt}}{W_{Pjt}}\right) S_{Tjt}}\right), \tag{18}$$

S_{Tjt} is the share of temporary (as opposed to permanent) workers in our data, Δ_{jt} is the gap between the wage of permanent workers W_{Pjt} and their shadow wage as defined in Equation (C.2), and $\left(\frac{W_{Tjt} - W_{Pjt}}{W_{Pjt}}\right)$ is the wage premium of temporary workers.

Comparing Equation (16) to Equation (4) shows that allowing for productivity differences and adjustment costs amounts to correcting the labour and wage variables l_{jt} and w_{jt} in our data by c_{1jt} and c_{2jt}. The difficulty in implementing this correction is that our data lacks some of the constituent parts of the corrections c_{1jt} and c_{2jt}.

First, we estimate the wage premium of temporary workers $\left(\frac{W_{Tjt} - W_{Pjt}}{W_{Pjt}}\right)$ by regressing the log of the wage w_{jt} on the share of temporary workers S_{Tjt}, the share of white (as opposed to blue) collar workers, and the shares of engineers and technicians (as opposed to unskilled workers).[29] In addition to these descriptors of the work force, we include time dummies, region dummies, product submarket dummies, and an array of other firm characteristics, namely an index of technological sophistication, an index of market dynamism, a dummy for the identification of

---

27. While we assume that the elasticity of substitution between permanent and temporary labour is unity, Aguirregabiria and Alonso-Borrego (2009) assume that is infinite. We leave it to future research to estimate the elasticity of substitution rather than assume it.

28. Allowing for non-convex adjustment costs requires structurally estimating the parameters of a dynamic model as in Aguirregabiria and Alonso-Borrego (2009).

29. Except for the share of temporary workers, we have these measures in the year a firm enters the sample and then every four years. In what follows we assume that the work force is unchanging in the interim. We further assume that the joint distribution (*e.g.* temporary white collar engineers) is the product of the marginal distributions that we observe.

ownership and control, and a third-order polynomial in age. We take the dummies and other firm characteristics to be exogenous to the firm, at least in the short run. The estimated coefficient $\widehat{\gamma}_T$ on S_{Tjt} is the wage premium of temporary workers $\left(\frac{W_{Tjt}-W_{Pjt}}{W_{Pjt}}\right)$. As Column (1) of Table 4 shows, $\widehat{\gamma}_T$ is negative and significant, in line with our expectation that temporary workers earn less than permanent workers. The wage regression fits the data well (Column (2)).

Second, because the first-order conditions for permanent and temporary labour presume an interior solution, we exclude observations with $S_{Tjt}=0$ and thus $L_{Tjt}=0$ from the analysis (compare Columns (3) and (4) of Table 4 to Columns (1) and (2) of Table 1). To estimate α and Δ_{jt} we note that the first-order conditions imply

$$\frac{W_{Tjt}S_{Tjt}}{W_{Pjt}(1-S_{Tjt})} = \left(\frac{1-\alpha}{\alpha}\right)(1+\Delta_{jt}). \tag{19}$$

Hence, if $E(\Delta_{jt})=0$ for a given firm, then using the definition of $\widehat{\gamma}_T$ and appealing to the analogy principle we obtain the firm-specific estimate

$$\left(\widehat{\frac{1-\alpha_j}{\alpha_j}}\right) = (1+\widehat{\gamma}_T)\frac{1}{T_j}\sum_{t=1}^{T_j}\frac{S_{Tjt}}{1-S_{Tjt}},$$

where T_j is the number of observations for firm j, and

$$\widehat{\Delta}_{jt} = \frac{(1+\widehat{\gamma}_T)}{\left(\widehat{\frac{1-\alpha_j}{\alpha_j}}\right)}\frac{S_{Tjt}}{1-S_{Tjt}} - 1.$$

These estimates, while not consistent for the small values of T_j in our data, ensure that our model is consistent with the share of temporary workers.[30] Column (5) of Table 4 shows that the average of the implied firm-specific estimate $\widehat{\alpha}$ ranges from 0.66 to 0.88 across industries. While $\widehat{\alpha}$ varies across the firms within an industry, it seems clear that the productivity of permanent workers differs from that of temporary workers.

With the corrections c_{1jt} and c_{2jt} in hand, we recover productivity from the aggregate labour demand function in Equation (16). We again exclude observations with $S_{Tjt}=0$ and thus $L_{Tjt}=0$ from the analysis because the aggregate labour demand function presumes an interior solution for permanent and temporary workers. The production function estimates in Column (6)–(8) of Table 4 are broadly comparable to those in Columns (4)–(6) of Table 2. The most visible changes are as expected in the labour coefficient, which increases in industries 1, 2, 3, 6, and 7 and decreases in industries 4, 8, 9, and 10. The capital coefficient decreases in industries 1, 2, 3, 4, and 6 and increases in industries 7, 8, and 10. The materials coefficient increases slightly in industries 2, 4, and 8 and remains essentially the same in the remaining industries. The absence of systematic changes suggests that combining the hours worked of permanent and temporary workers in our leading specification does not distort the estimates in a major way. Note, however, that at a 5% level we reject the validity of the moment conditions in industry 7 and that we have been unable to obtain second-step GMM estimates in industry 9 (Columns (9) and (10)).

Our conclusions regarding the link between R&D and productivity are similarly robust (compare the average difference in expected productivity between performers and non-performers in Column (10) of Table 4 to Column (1) of Table 6). We return to productivity levels and growth below.

---

30. Decomposing the right-hand side of Equation (19) into firm and time fixed effects to estimate α and Δ_{jt} yields similar results.

TABLE 4

Production function estimates, specification tests, and productivity levels: permanent and temporary workers

Industry	Wage regression[a]		Obs.[b]	Firms[b]	$\widehat{\alpha}$[b]	GMM[a,b]			Overidentif. restr. test[b]		Diff. of means[b]
	Temp. (s. e.)	R^2			Mean (s. d.)	Cap. (s. e.)	Lab. (s. e.)	Mat. (s. e.)	$\chi^2(df)$	p val.	≤ 200 > 200
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. Metals and metal products	−0.405 (0.063)	0.599	1046	262	0.786 (0.200)	0.094 (0.009)	0.173 (0.015)	0.692 (0.012)	67.916 (51)	0.057	0.010 0.003
2. Non-metallic minerals	−0.414 (0.105)	0.701	477	129	0.827 (0.169)	0.067 (0.012)	0.209 (0.014)	0.706 (0.012)	53.832 (47)	0.229	0.050 0.045
3. Chemical products	−0.494 (0.086)	0.680	1019	250	0.852 (0.165)	0.055 (0.010)	0.196 (0.016)	0.741 (0.012)	55.434 (47)	0.187	0.025 0.041
4. Agric. and ind. machinery	−0.183 (0.088)	0.594	439	118	0.817 (0.188)	0.046 (0.015)	0.247 (0.020)	0.691 (0.013)	44.974 (44)	0.431	−0.060 0.024
6. Transport equipment	−0.625 (0.072)	0.664	538	138	0.884 (0.135)	0.079 (0.012)	0.208 (0.018)	0.673 (0.016)	42.877 (47)	0.644	0.040 0.065
7. Food, drink, and tobacco	−0.366 (0.064)	0.659	1222	290	0.732 (0.230)	0.081 (0.011)	0.159 (0.015)	0.756 (0.011)	67.638 (46)	0.021	−0.011 0.015
8. Textile, leather, and shoes	−0.323 (0.055)	0.620	1023	269	0.703 (0.260)	0.065 (0.011)	0.240 (0.017)	0.640 (0.014)	61.745 (51)	0.144	0.008 0.016
9. Timber and furniture[c]	−0.323 (0.070)	0.618	499	129	0.662 (0.222)	0.128 (0.028)	0.121 (0.031)	0.728 (0.030)	– –	–	−0.060 0.060
10. Paper and printing products	−0.259 (0.107)	0.694	518	142	0.805 (0.182)	0.131 (0.015)	0.223 (0.022)	0.624 (0.019)	44.424 (42)	0.370	−0.054 0.047

[a] All standard errors are robust to heteroskedasticity and autocorrelation.
[b] Conditional on nonzero temporary workers in period $t-1$.
[c] First-step GMM estimates.

5.2. *Non-linearity and uncertainty*

A key contribution of this article is to endogenize the productivity process. We assess the role of R&D by comparing the controlled Markov process in Equation (3) with an exogenous Markov process (as in OP, LP, and ACF). To this end, we test whether R&D can be excluded from the conditional expectation function $g(\cdot)$ so that $g_{00} = g_{10}$ and $g_{01}(h_{ljt-1} - \lambda_l) = g_{11}(h_{ljt-1} - \lambda_l, r_{jt-1})$ in Equation (13). As Columns (1) and (2) of Table 5 show, in all industries the exogenous Markov process is clearly rejected.

We next ask if the evolution of productivity depends on the amount of R&D expenditures in addition to whether or not a firm engages in R&D. The data reject the restriction $g_{11}(h_{ljt-1} - \lambda_l, r_{jt-1}) = g_{11}(h_{ljt-1} - \lambda_l)$ in all industries at a 5% level.

Finally we ask if further lags of R&D expenditures matter. The impact of r_{jt-1} on ω_{jt} in Equation (3) may be moderated by r_{jt-2}. Or it may be that R&D expenditures take more than one period to come to fruition, although the available evidence points to rather short gestation periods (see pp. 82–84 of Pakes and Schankerman (1984) and the references therein). For the four industries with the largest number of observations we replace $g_{11}(h_{ljt-1} - \lambda_l, r_{jt-1})$ in Equation (13) by $g_{11}(h_{ljt-1} - \lambda_l, r_{jt-1}, r_{jt-2})$, where we set $r_{jt-2} = 0$ if $R_{jt-2} = 0$. We reject our leading specification at a 5% level in industry 3 but not in industries 1, 7, and 8. In industry 3 (chemical products) investments in knowledge may be especially cumulative in that many research projects continue—and directly build on—previous ones. This may also create an incentive for firms "to keep R&D going", and indeed industry 3 has the largest share of stable performers (see again Column (11) of Table 1). Overall, however, we feel that our leading specification is a reasonable first pass at endogenizing the productivity process.

Non-linearity: As a step toward exploring the link between R&D and productivity, we test whether the conditional expectation function $g(\cdot)$ is separable in already attained productivity ω_{jt-1} and R&D expenditures r_{jt-1} so that $g_{11}(h_{ljt-1} - \lambda_l, r_{jt-1}) = g_{111}(h_{jt-1} - \lambda_l) + g_{112}(r_{jt-1})$ in Equation (13). Columns (3) and (4) of Table 5 indicate that this restriction is rejected in eight industries at a 5% level and in all industries at a 10% level. Hence, the R&D process can hardly be considered separable. From the economic point of view this stresses that the impact of current R&D on future productivity depends crucially on current productivity, and that current and past investments in knowledge interact in a complex fashion.

To illustrate the nature of these interactions, we compute the percentage of observations where $\frac{\partial^2 g(\omega_{jt-1}, r_{jt-1})}{\partial \omega_{jt-1} \partial R_{jt-1}} = \frac{1}{R_{jt-1}} \frac{\partial^2 g(\omega_{jt-1}, r_{jt-1})}{\partial \omega_{jt-1} \partial r_{jt-1}}$ is significantly positive (negative) at a 5% level so that productivity and (the level of) R&D expenditures are, at least locally, complements (substitutes) in the accumulation of productivity. While in industry 8 productivity and R&D expenditures are substitutes for a substantial fraction of observations, there is evidence of complementarities in industries 1, 2, 3, 4, 6, 9, and 10. Current productivity thus tends to reinforce the impact of current R&D on future productivity.

Uncertainty: Turning from non-linearity to uncertainty, Column (5) of Table 5 tells us the ratio of the variance of the random shock e_{jt} to the variance of unobserved productivity ω_{jt}. Despite differences among industries, the variances are quite similar in magnitude. This suggests that unobserved productivity is at least as important in explaining the data as the host of other factors that are embedded in the random shock.

Column (6) of Table 5 gives the ratio of the variance of the productivity innovation ξ_{jt} to the variance of actual productivity ω_{jt}. The ratio shows that the unpredictable component accounts for a large part of productivity, between 25% and 75%. Interestingly enough, a high degree of uncertainty in the R&D process seems to be characteristic for both some of the most and some of

TABLE 5

Non-linearity and uncertainty and comparison to knowledge capital model

Industry	Exogeneity test		Separability test		$\frac{Var(e_{it})}{Var(\omega_{it})}$	$\frac{Var(\xi_{it})}{Var(\omega_{it})}$	Knowledge capital model tests					
							Basic		Generalization 1		Generalization 2	
	$\chi^2(10)$	p val.	$\chi^2(3)$	p val.			$N(0,1)$	p val.	$N(0,1)$	p val.	$N(0,1)$	p val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1. Metals and metal products	65.55	0.000	16.360	0.001	0.735	0.407	−2.815	0.002	−2.431	0.008	−1.987	0.023
2. Non-metallic minerals	92.65	0.000	13.027	0.005	0.842	0.410	−2.041	0.021	−1.541	0.062	−0.784	0.216
3. Chemical products	40.79	0.000	8.647	0.034	0.749	0.244	−3.239	0.001	−2.090	0.018	−1.400	0.081
4. Agric. and ind. machinery	51.88	0.000	11.605	0.009	1.410	0.505	−2.693	0.004	−1.588	0.056	−1.493	0.068
6. Transport equipment	56.85	0.000	18.940	0.000	1.626	0.524	−2.317	0.010	−2.042	0.021	−1.821	0.034
7. Food, drink, and tobacco	38.29	0.000	7.186	0.066	1.526	0.300	−3.263	0.001	−2.499	0.006	−0.901	0.184
8. Textile, leather, and shoes	29.91	0.001	18.417	0.000	1.121	0.750	−2.770	0.003	−1.788	0.037	−1.488	0.068
9. Timber and furniture	118.17	0.000	32.260	0.000	1.417	0.515	−2.510	0.006	−2.097	0.018	−1.028	0.152
10. Paper and printing products	59.73	0.000	23.249	0.000	0.713	0.433	−3.076	0.001	−2.210	0.014	−1.595	0.055

the least R&D intensive industries. We come back to the economic significance of the uncertainties inherent in the R&D process in Section 6.4.

5.3. *Comparison to knowledge capital model*

The knowledge capital model of Griliches (1979) has remained a cornerstone of the productivity literature. It augments a standard production function with a measure of the current state of technical knowledge yielding

$$y_{jt} = \beta_0 + \beta_t t + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \varepsilon c_{jt} + e_{jt}, \tag{20}$$

where c_{jt} is the stock of knowledge capital of firm j in period t and ε is the elasticity of output with respect to this stock.

While the knowledge capital model has been used in hundreds of studies on firm-level productivity, the underlying empirical strategy has changed little over the years (see the surveys by Mairesse and Sassenou (1991), Griliches (1995, 2000), and Hall *et al.* (2010)). Almost all studies use a simple perpetual inventory or declining balance methodology to construct the stock of knowledge capital from the firm's observed R&D expenditures as $C_{jt} = (1-\delta)C_{jt-1} + R_{jt-1}$, where δ is a single constant rate of depreciation. This assumes linear and certain accumulation of knowledge from period to period in proportion to R&D expenditures as well as linear and certain depreciation.

From a practical point of view, there are at least two problems. First, estimating the rate of depreciation δ using distributed lag models is notoriously difficult, even in case of physical capital (Pakes and Schankerman, 1984; Pakes and Griliches, 1984; Nadiri and Prucha, 1996). Because our own attempts at estimating δ together with the parameters in Equation (20) have largely failed, we follow Hall and Mairesse (1995) and assume a rate of depreciation of 0.15 (see also p. 16 of Hall *et al.*, 2010). Second, the available history of R&D expenditures is often not very long and there rarely is information on the initial stock of knowledge capital. We again follow Hall and Mairesse (1995) and estimate the initial condition from the date of birth of the firm by extrapolating its average R&D expenditures during the time that it is observed.

We test this basic—albeit most widely used—form of the knowledge capital model against our dynamic investment model.[31] The non-nested test very clearly rejects the knowledge capital model in its basic form, see Columns (7) and (8) of Table 5.

Replacing εc_{jt} by ω_{jt} in Equation (20) shows that the basic knowledge capital model can be understood as a special case of our model in which the stochastic process that governs productivity has degenerated to a deterministic process with a particular functional form for $g(\cdot)$. Given that we have already shown that non-linearity and uncertainty play a large role in the link between R&D and productivity, it is therefore not surprising that the data favour our model.

There are few attempts to relax the linearity and certainty assumptions of the basic knowledge capital model. Griliches (1979) allows for a random shock to the constructed stock of knowledge capital. Being entirely transitory, however, this shock is absorbed by the error term of the production function and hence does not capture the uncertainties inherent in the R&D process (p. 100).

---

31. For firms without a positive stock of knowledge capital we drop the term εc_{jt} from Equation (20) and specify a different constant.

Hall and Hayashi (1989) and Klette (1996) allow for non-linearities in the form of complementarities and uncertainties in the accumulation of knowledge. To prevent the stock of knowledge capital from vanishing if a firm does not engage in R&D in a single year the law of motion is $C_{jt} = C_{jt-1}^{\sigma}(1+R_{jt-1})^{1-\sigma}e^{\frac{1}{\varepsilon}\xi_{jt}}$. Because of the random shock ξ_{jt} it is no longer possible to construct the stock of knowledge capital from the firm's observed R&D expenditures, contrary to the standard approach in the knowledge capital literature. To be able to proceed without constructing the stock of knowledge capital Hall and Hayashi (1989) and Klette (1996) develop variance-components and pseudo-difference approaches to estimation.

We recast their law of motion in our setting by taking logs and letting $\omega_{jt} = \varepsilon c_{jt}$ to write

$$\omega_{jt} = \sigma \omega_{jt-1} + \varepsilon(1-\sigma)\ln(1+\exp(r_{jt-1})) + \xi_{jt}$$

and hence $\omega_{jt} = g(\omega_{jt-1}, r_{jt-1}) + \xi_{jt}$, a special case of our controlled first-order Markov process with a particular functional form for $g(\cdot)$. The non-nested test rejects this first generalization of the knowledge capital model in seven industries at a 5% level and in all industries at a 10% level, see Columns (9) and (10) of Table 5.[32]

Aside from the additional flexibility from non-parametrically estimating the conditional expectation function $g(\cdot)$, our setting differs from Hall and Hayashi (1989) and Klette (1996) in that it allows us to infer a firm's unobserved productivity from its observed decisions. In contrast, neither Hall and Hayashi (1989) nor Klette (1996) can recover a firm's stock of knowledge capital or assess its productivity even after the model has been estimated.

Our second generalization builds on Griliches and Mairesse (1998) and combines the knowledge capital model with the literature following OP (see pp. 190–194 and also pp. 276–278 of Griliches (1998) and pp. 57–58 of Griliches (2000)). It endogenizes productivity by modelling it as $\varepsilon c_{jt} + \omega_{jt}$, where $C_{jt} = (1-\delta)C_{jt-1} + R_{jt-1}$ is constructed from the firm's observed R&D expenditures (as in the knowledge capital literature) and ω_{jt} follows an exogenous Markov process (as in OP, LP, and ACF). Hence, Equation (20) becomes

$$y_{jt} = \beta_0 + \beta_t t + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \varepsilon c_{jt} + \omega_{jt} + e_{jt}.$$

This model has two Markov processes, one for the deterministic component of productivity (εc_{jt}) and one for the stochastic component (ω_{jt}).[33] Because the sum of two Markov processes is not necessarily a Markov process, it is not nested in our dynamic investment model with just one Markov process. In this sense our model is more parsimonious. The non-nested test nevertheless rejects this second generalization of the knowledge capital model in two industries at a 5% level and in six industries at a 10% level, see Column (11) and (12) of Table 5.

In sum, our dynamic investment model can be viewed either as a generalization of the basic knowledge capital model or as a practical alternative to generalizations of the knowledge capital model. The unspecified form of the law of motion and the random nature of accumulation are important advantages of our setting. Moreover, by treating productivity as unobserved and inferring it from a firm's observed decisions, we circumvent the initial condition problem

---

32. We obtain very similar results if we start with a law of motion of the form $C_{jt} = (1-\delta)C_{jt-1} + R_{jt-1} + \frac{1}{\varepsilon}C_{jt-1}\xi_{jt} = C_{jt-1}\left(1-\delta+\frac{R_{jt-1}}{C_{jt-1}}+\frac{1}{\varepsilon}\xi_{jt}\right)$ so that the effect of the rate of investment in knowledge $\frac{R_{jt-1}}{C_{jt-1}}$ has an unpredictable component $\frac{1}{\varepsilon}\xi_{jt}$. Taking logs and letting $\omega_{jt} = \varepsilon c_{jt}$, this law of motion can be written as $\omega_{jt} \simeq \omega_{jt-1} + \varepsilon\left(\frac{\exp(r_{jt-1})}{\exp(\omega_{jt-1}/\varepsilon)}-\delta\right)+\xi_{jt}$.

33. Allowing for a random shock in the law of motion for the stock of knowledge capital as in Hall and Hayashi (1989) and Klette (1996) leads to a model with two non-degenerate Markov processes that is beyond the scope of this article. A difficulty arises because the first-order conditions for static inputs cannot identify εc_{jt} separately from ω_{jt}.

in the basic knowledge capital model along with the practical problem of estimating a rate of depreciation for the stock of knowledge capital. Relative to the estimation approaches in Hall and Hayashi (1989) and Klette (1996), inferring a firm's unobserved productivity from its observed decisions gives us the ability to assess the role of R&D in determining the differences in productivity across firms and the evolution of firm-level productivity over time.

6. R&D AND PRODUCTIVITY

To assess the role of R&D in determining the differences in productivity across firms and the evolution of firm-level productivity over time, we examine five aspects of the link between R&D and productivity in more detail: productivity levels and growth, the return to R&D, the persistence in productivity, and the rate of return.

6.1. *Productivity levels*

To describe differences in expected productivity $g(\omega_{jt-1}, r_{jt-1})$ between firms that perform R&D and firms that do not, we employ kernels to estimate the distribution functions for the subsamples of observations with and without R&D. To be able to interpret these and other descriptive measures in the remainder of the article as representative aggregates, we replicate the subsample of small firms $\frac{70}{5} = 14$ times before pooling it with the subsample of large firms.[34] Figure 1 shows the distribution functions for performers (solid line) and non-performers (dashed line) for each industry. In most industries the distribution function for performers is to the right of the distribution function for non-performers. This strongly suggests stochastic dominance. In contrast, the distribution functions cross in industries 2 and 4 that exhibit medium and high innovative activity as well as in industries 9 and 10 that exhibit low innovative activity.

Mean: Before more formally comparing the distributions themselves, we compute the difference in means as

$$\bar{g}_0 - \bar{g}_1 = \frac{1}{NT_0} \sum_j \sum_t 1(R_{jt-1} = 0) \widehat{g}_{01}\left(\widehat{h}_{ljt-1}\right)$$
$$- \frac{1}{NT_1} \sum_j \sum_t 1(R_{jt-1} > 0) \left[\left(\widehat{g_{10} - g_{00}}\right) + \widehat{g}_{11}\left(\widehat{h}_{ljt-1}, r_{t-1}\right)\right],$$

where NT_0 and NT_1 are the size of the subsamples of observations without and with R&D, respectively. Then the test statistic

$$t = \frac{\bar{g}_0 - \bar{g}_1}{\sqrt{Var\left(\widehat{g}_{01}\left(\widehat{h}_{ljt-1}\right)\right)/(NT_0 - 1) + Var\left(\widehat{g}_{11}\left(\widehat{h}_{ljt-1}, r_{jt-1}\right)\right)/(NT_1 - 1)}}$$

follows a t distribution with $\min(NT_0, NT_1) - 1$ degrees of freedom. To account for the survey design we conduct separate tests for the subsamples of small and large firms.

---

34. Because replicating the subsample of small firms distorts variances in the pooled sample, we have been careful to use replication only if we compute averages. Whenever we compute variances we do so either separately for small and large firms or we pool small and large firms but without first replicating the small firms.
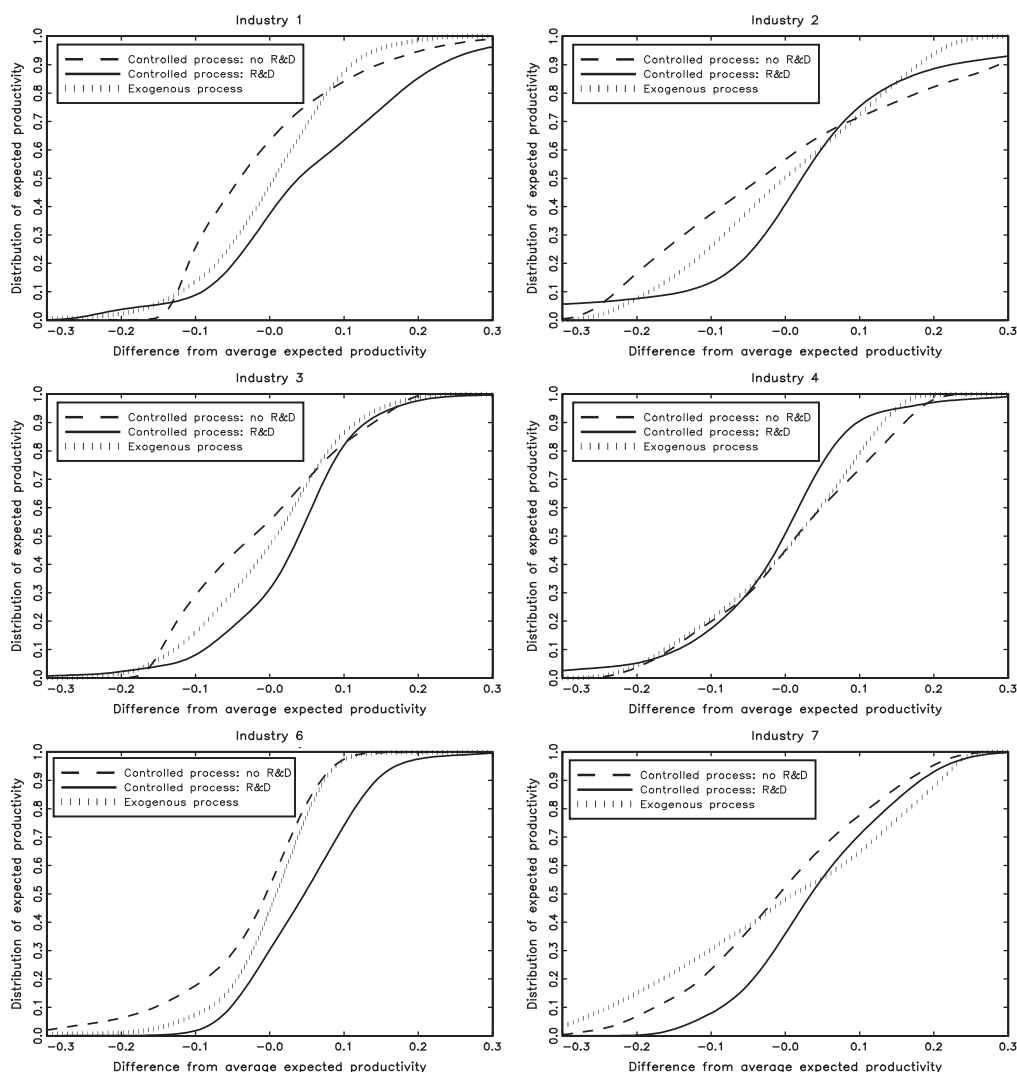
FIGURE 1

Distribution of expected productivity

Column (1) of Table 6 reports the difference in means $\bar{g}_1 - \bar{g}_0$ (rather than $\bar{g}_0 - \bar{g}_1$ to aid intuition). The difference in means is positive for firms of all sizes in all industries that exhibit medium or high innovative activity, with the striking exception of industry 4.[35] The differences are sizable, typically between 3% and 5%. They are often larger for the smaller firms. In the two industries that exhibit low innovative activity, however, the difference in means is negative for small firms. Nevertheless, as can be seen from Columns (2) and (3), at a 5% level the test rejects the hypothesis that the mean of expected productivity is higher for performers than non-performers solely in industry 4.

---

35. A possible explanation for the abnormal result in industry 4 is the considerable heterogeneity in activities across the firms within an industry that arises due the level of aggregation we use.
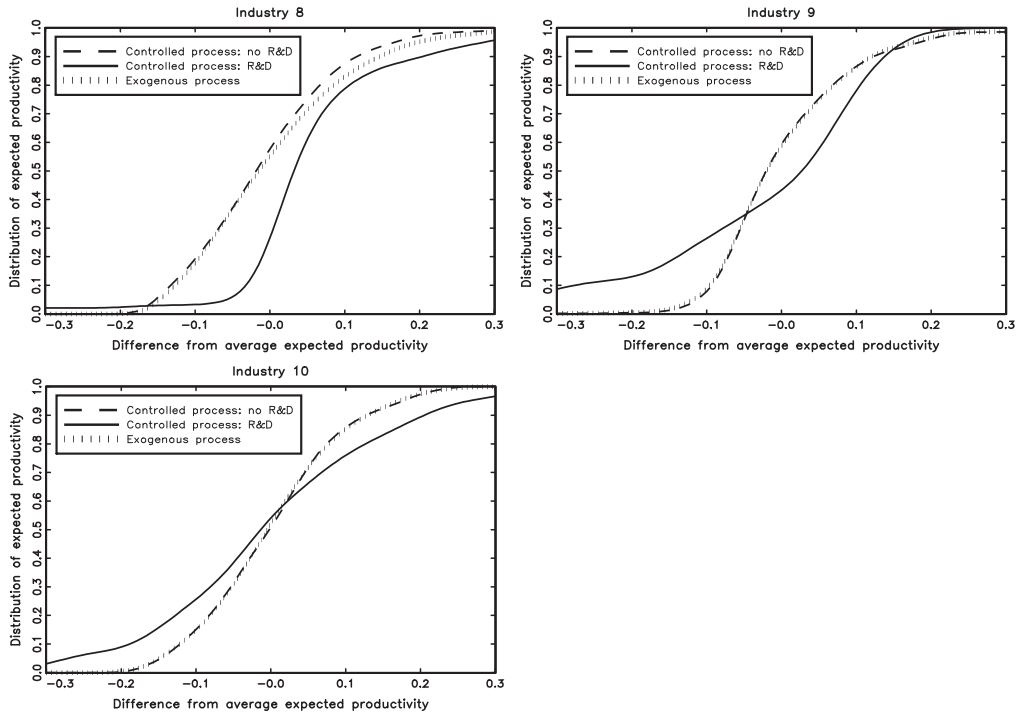
FIGURE 1

Continued

Distribution: To compare the distribution themselves, we apply a Kolmogorov–Smirnov test (see Barrett and Donald (2003) and Delgado *et al.* (2002) for similar applications). Since this test requires that the observations in each sample are independent, we consider as the variable of interest the average of expected productivity for each firm. For occasional performers we average only over the years with R&D (and discard the years without R&D).

Let $G_{N_1}(\cdot)$ and $F_{N_0}(\cdot)$ be the empirical distribution functions of performers and non-performers, respectively, with N_1 and N_0 being the number of performers and non-performers. We apply a two-sided test of the hypothesis $G_{N_1}(\overline{g}) - F_{N_0}(\overline{g}) = 0$ for all $\overline{g}$, *i.e.* the distributions $G_{N_1}(\cdot)$ and $F_{N_0}(\cdot)$ of expected productivity are equal, and a one-sided test of the hypothesis $G_{N_1}(\overline{g}) - F_{N_0}(\overline{g}) \leq 0$ for all $\overline{g}$, *i.e.* the distribution $G_{N_1}(\cdot)$ of expected productivity of performers stochastically dominates the distribution $F_{N_0}(\cdot)$ of expected productivity of non-performers. The test statistics are

$$S^1 = \sqrt{\frac{N_0 N_1}{N_0 + N_1}} \max_{\overline{g}} \left\{ |G_{N_1}(\overline{g}) - F_{N_0}(\overline{g})| \right\}, \quad S^2 = \sqrt{\frac{N_0 N_1}{N_0 + N_1}} \max_{\overline{g}} \left\{ G_{N_1}(\overline{g}) - F_{N_0}(\overline{g}) \right\},$$

respectively, and the probability values can be computed using the limiting distributions $P(S^1 > c) = -2 \sum_{k=1}^{\infty} (-1)^k \exp(-2k^2 c^2)$ and $P(S^2 > c) = \exp(-2c^2)$.

Because the tests tend to be inconclusive when the number of firms is small, we limit them to cases in which we have at least 20 performers and 20 non-performers. As can be seen from

TABLE 6
Productivity levels

Industry	Size	Diff. of means	Mean with R&D is greater		Kolmogorov–Smirnov tests[a]			
					Distributions are equal		Distribution with R&D dominates	
			t	p val.	S_1	p val.	S_2	p val.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Metals and metal products	≤ 200	0.051	−4.992	1.000	1.720	0.005	0.388	0.740
	> 200	0.038	−2.920	0.998				
2. Non-metallic minerals	≤ 200	0.030	−1.320	0.905	1.167	0.131	0.938	0.172
	> 200	0.025	−1.313	0.904				
3. Chemical products	≤ 200	0.033	−4.531	1.000	1.440	0.032	0.351	0.782
	> 200	0.028	−2.003	0.975				
4. Agric. and ind. machinery	≤ 200	−0.023	1.807	0.036	0.988	0.283	0.988	0.142
	> 200	−0.077	3.644	0.000				
6. Transport equipment	≤ 200	0.073	−6.377	1.000	1.770	0.004	0.000	1.000
	> 200	0.038	−4.126	1.000				
7. Food, drink, and tobacco	≤ 200	0.034	−3.476	1.000	1.708	0.006	0.198	0.925
	> 200	0.028	−3.073	0.999	0.610	0.850	0.227	0.902
8. Textile, leather, and shoes	≤ 200	0.044	−2.543	0.994	3.173	0.000	0.335	0.799
	> 200	0.022	−1.552	0.939	0.511	0.957	0.426	0.696
9. Timber and furniture	≤ 200	−0.020	0.724	0.237				
	> 200	0.014	−0.610	0.721				
10. Paper and printing products	≤ 200	−0.013	0.590	0.279	0.930	0.353	0.709	0.366
	> 200	0.028	−1.400	0.917				

[a] Applied to a firm's average expected productivity when each sample has more than 20 firms.

Columns (4)–(7) of Table 6 we reject equality of the distributions for performers and non-performers in five of ten cases at a 5% level. We cannot reject stochastic dominance anywhere at a 5% level.[36,37]

Omitting R&D expenditures: To illustrate the consequences of omitting R&D expenditures from the Markov process of unobserved productivity, we have added the so-obtained distribution functions to Figure 1 (dotted line). Comparing them to the distribution functions for the controlled Markov process in our leading specification reveals that the exogenous process amounts to averaging over firms with distinct innovative activities and hence blurs important differences in the impact of the investment in knowledge on the productivity of firms.

Redoing the above tests for the exogenous Markov process yields striking results: we no longer reject equality of the distributions of expected productivity of performers and non-performers in three of the five cases where we rejected for a controlled Markov process. Figure 2 shows at the example of industry 6 that the distribution functions for performers (solid line) and non-performers (dashed line) are virtually indistinguishable if an exogenous Markov process is assumed. This once more makes apparent that omitting R&D expenditures substantially distorts the retrieved

36. Treating occasional performers as non-performers by averaging only over the years without R&D (and discarding the years with R&D), we reject equality of the distributions in three cases at a 5% level and in five cases at a 10% level. We cannot reject stochastic dominance anywhere at a 5% level.

37. Linton *et al.* (2005) relax the independence assumption in Barrett and Donald (2003) and their test for stochastic dominance can be applied directly to the expected productivity of each firm in each period (rather than an average thereof). The test statistic is the same as in Barrett and Donald (2003), but the critical value has to be computed by a subsampling method. Applying this alternative test to the 12 cases in which we have at least 80 observations with R&D and 80 observations without R&D, we cannot reject stochastic dominance anywhere at a 5% level.
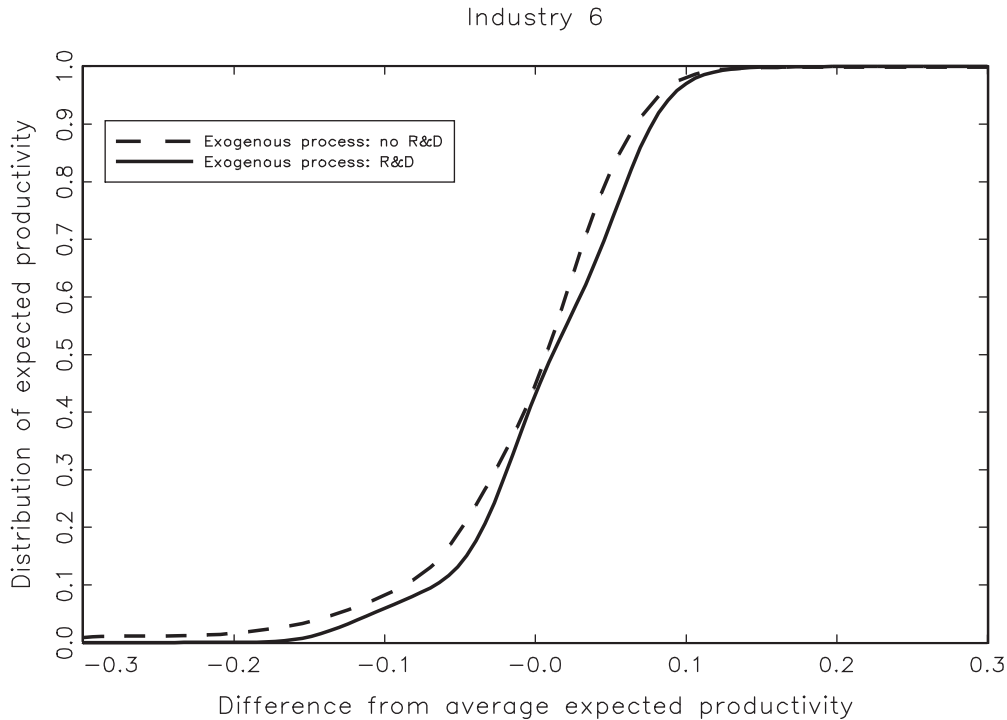
Industry 6

FIGURE 2

Distribution of expected productivity. Exogenous Markov process

productivities and cautions against the popular approach of first estimating productivity from a model with an exogenous Markov process and then regressing estimated productivity on its determinants such as R&D expenditures or export market participation.

6.2. *Return to R&D and persistence in productivity*

How hard must a firm work to maintain and advance its productivity? Since a change in the conditional expectation function $g(\cdot)$ can be interpreted as the expected percentage change in total factor productivity, $\frac{\partial g(\omega_{jt-1}, r_{jt-1})}{\partial r_{jt-1}}$ is the elasticity of output with respect to R&D expenditures or a measure of the return to R&D.[38] Similarly, $\frac{\partial g(\omega_{jt-1}, r_{jt-1})}{\partial \omega_{jt-1}}$ is the elasticity of output with respect to already attained productivity. $\frac{\partial g(\omega_{jt-1}, r_{jt-1})}{\partial \omega_{jt-1}}$ is the degree of persistence in the productivity process or a measure of inertia. It tells us the fraction of past productivity that is carried forward into current productivity. Note that the elasticities of output with respect to R&D expenditures and already attained productivity vary from firm to firm with already attained productivity and R&D expenditures. Our model thus allows us to recover the distribution of these elasticities and to describe the heterogeneity across firms.

---

38. If we consider a *ceteris paribus* increase in R&D expenditures that changes ω_{jt} to $\tilde{\omega}_{jt}$, then $\tilde{\omega}_{jt} - \omega_{jt}$ approximates the effect of this change in productivity on output in percentage terms, *i.e.* $(\tilde{Y}_{jt} - Y_{jt})/Y_{jt} = \exp(\tilde{\omega}_{jt} - \omega_{jt}) - 1 \simeq \tilde{\omega}_{jt} - \omega_{jt}$. That is, the change in ω_{jt} shifts the production function and hence measures the change in total factor productivity. Also $g(\cdot)$ and ξ_{jt} can be interpreted in percentage terms and decompose the change in total factor productivity.

Return to R&D: Columns (1)–(4) of Table 7 present the quartiles of the distribution of the elasticity with respect to R&D expenditures along with a weighted average computed as $\frac{1}{T} \sum_t \sum_j \mu_{jt} \frac{\partial \widehat{g}(\widehat{h}_{ljt-1}, r_{jt-1})}{\partial r_{jt-1}}$, where the weights $\mu_{jt} = Y_{jt}/\sum_j Y_{jt}$ are given by the share of output of a firm. There is a considerable amount of variation across industries and the firms within an industry. The returns to R&D at the first, second, and third quartile range between -0.036 and 0.009, -0.012 and 0.022, and 0.000 and 0.051, respectively. Their average is close to 0.015, varying from -0.006 to 0.046 across industries.

Negative returns to R&D are legitimate and meaningful in our setting, although some of them may be an artefact of the non-parametric estimation of $g(\cdot)$ at the boundaries of the support. A negative return at the margin is consistent with an overall positive impact of R&D expenditures on output. A firm may invest in R&D to the point of driving returns below zero for a number of reasons including indivisibilities and strategic considerations such as a loss of an early-mover advantage. This type of effect is excluded by the functional form restrictions of the knowledge capital model, in particular the assumption that the stock of knowledge capital depreciates at a constant rate. More generally, it is plausible that investments in knowledge take place in response to existing knowledge becoming obsolete or *vice versa* that investments render existing knowledge obsolete. Our model captures this interplay between adding "new" knowledge and keeping "old" knowledge.

Degree of persistence: The degree of persistence can be computed separately for performers using the conditional expectation function $g_1(\cdot)$ that depends both on already attained productivity and R&D expenditures and for non-performers using $g_0(\cdot)$ that depends solely on already attained productivity. Columns (5)–(10) of Table 7 summarize the distributions for performers and non-performers.

Again there is a considerable amount of variation across industries and the firms within an industry. Nevertheless, non-performers tend to enjoy a higher degree of persistence than performers in industries 1, 2, 3, 4, 7, and 8. An intuitive explanation for this finding is that non-performers learn from performers, but by the time this happens the transferred knowledge is already entrenched in the industry and therefore more persistent. Put differently, common practice may be "stickier" than best practice.

The degree of persistence for performers is negatively related to the degree of uncertainty in the productivity process as measured by the ratio of the variance of the productivity innovation ξ_{jt} to the variance of actual productivity ω_{jt}. That is, productivity is less persistent in an industry where a large part of its variance is due to random shocks that represent the uncertainties inherent in the R&D process. Figure 3 illustrates this relationship between persistence and uncertainty at the level of the industry.

Comparison to knowledge capital model: To facilitate the comparison with the existing literature, we have estimated the knowledge capital model as given in Equation (20).[39] Column (11) of Table 7 presents the estimate of the elasticity of output with respect to the stock of knowledge capital from the knowledge capital model. In addition to the gross-output version in Equation (20) we have also estimated a value-added version of the knowledge capital model (Column (13)). In contrast to our model, the knowledge capital model yields one number—an average elasticity—per industry. The elasticity of output with respect to the stock of knowledge

---

39. To improve the estimates we impose the widely accepted constraint of constant returns to scale in the conventional inputs.

TABLE 7
Elasticities of output with respect to R&D expenditures and already attained productivity

Industry	Elasticity wrt. R_{jt-1} [a]				Elasticity wrt. ω_{jt-1} [b]						Knowledge capital model Elasticity wrt. C_{jt} and R_{jt-1} [c]			
					Performers			Non-performers			Gross output		Value added	
	Q_1	Q_2	Q_3	Mean	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3	ε (s. e.)	Mean	ε (s. e.)	Mean
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Metals and metal products	−0.013	0.007	0.021	0.022	0.504	0.619	0.755	0.441	0.759	0.901	0.003 (0.005)	0.001	0.025 (0.016)	0.005
2. Non-metallic minerals	−0.018	−0.012	0.000	−0.006	0.433	0.477	0.575	0.377	0.646	0.878	0.010 (0.006)	0.002	0.035 (0.017)	0.008
3. Chemical products	0.009	0.011	0.014	0.013	0.459	0.523	0.634	0.547	0.815	0.947	0.018 (0.004)	0.003	0.075 (0.011)	0.014
4. Agric. and ind. machinery	−0.017	−0.009	0.021	0.005	0.434	0.721	0.791	0.729	0.894	0.979	−0.003 (0.008)	−0.001	0.025 (0.016)	0.005
6. Transport equipment	−0.034	−0.008	0.010	0.020	0.404	0.615	0.727	0.423	0.513	0.646	0.009 (0.005)	0.002	0.017 (0.017)	0.004
7. Food, drink, and tobacco	−0.008	0.010	0.026	0.020	0.445	0.705	0.867	0.822	0.930	0.965	0.002 (0.006)	0.001	0.046 (0.012)	0.011
8. Textile, leather, and shoes	−0.003	0.014	0.051	0.046	0.090	0.325	0.626	0.491	0.605	0.689	0.011 (0.008)	0.002	0.018 (0.018)	0.003
9. Timber and furniture	−0.031	0.005	0.048	0.004	0.458	0.585	0.814	0.303	0.430	0.641	0.018 (0.011)	0.005	0.074 (0.030)	0.019
10. Paper and printing products	−0.036	0.022	0.049	0.013	0.405	0.676	0.812	0.569	0.644	0.670	0.013 (0.009)	0.004	0.041 (0.028)	0.012

[a] We trim 2.5% of observations at each tail of the distribution of h_{jt-1}.
[b] We trim observations below zero and above unity.
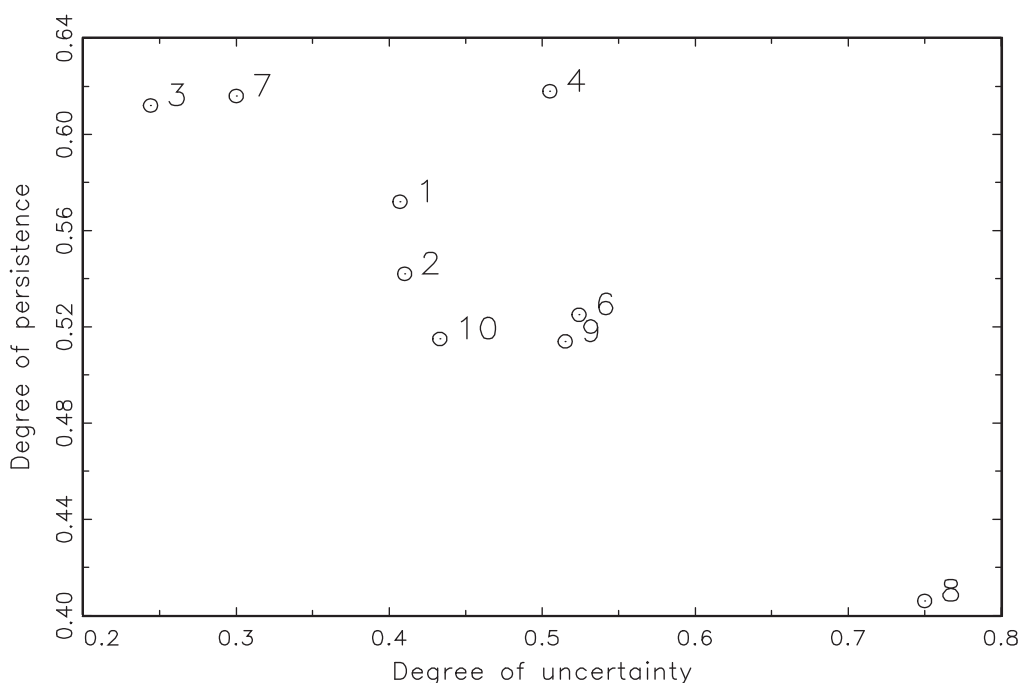[c] All standard errors are robust to heteroskedasticity and autocorrelation.

FIGURE 3

Persistence and uncertainty

capital tends to be small and rarely significant in the gross-output version but becomes larger in the value-added version.[40]

To convert the elasticity with respect to the stock of knowledge capital into an elasticity with respect to R&D expenditures that is comparable to our model, we multiply the former by R_{jt-1}/C_{jt}. Columns (12) and (14) of Table 7 show a weighted average of the so-obtained elasticities, where the weights $\mu_{jt} = Y_{jt}/\sum_j Y_{jt}$ are given by the share of output of a firm. The elasticities with respect to R&D expenditures from our model are higher than the highest elasticities from the knowledge capital model in five industries and lower but very close in two more industries. In addition, the elasticities obtained with our model have a non-normal, fairly spread out distribution. This sharply contrasts with the fact that the dispersion of elasticities in the knowledge capital model is purely driven by the distribution of the ratio R_{jt-1}/C_{jt} (since, recall, the knowledge capital model yields just an average of the elasticity with respect to the stock of knowledge capital).

Turning to persistence in productivity, note that the degree of persistence is $1-0.15=0.85$ by assumption in the knowledge capital model. In contrast, the degree of persistence in our model is much lower, in line with previous evidence from patent renewal decisions (see, *e.g.* Pakes and Schankerman, 1984; Schankerman and Pakes, 1986).

In sum, it appears that old knowledge is hard to keep but new knowledge is easy to add. Productivity is therefore considerably more fluid than what the knowledge capital literature suggests. As a consequence, a firm's position in the productivity distribution is considerably less stable than what the knowledge capital literature suggests, especially because a firm is repeatedly

---

40. Beneito (2001) and Ornaghi (2006) estimate comparable elasticities ranging from 0.04 to 0.10.

subjected to shocks that may make it hard for it to "break away" from its rivals and remain at or near the top of the productivity distribution.

6.3.   *Productivity growth*

We explore productivity growth from the point of view of what a firm expects when it makes its decisions in period $t-1$. Because ω_{jt-1} is known to the firm at the time it decides on r_{jt-1}, the expectation of productivity growth including the trend is

$$\beta_t + E[\omega_{jt} - \omega_{jt-1}|\omega_{jt-1}, r_{jt-1}] = \beta_t + g(\omega_{jt-1}, r_{jt-1}) - \omega_{jt-1}. \tag{21}$$

We estimate the average of the expectation of productivity growth as $\widehat{\beta}_t + \frac{1}{T}\sum_t\sum_j\mu_{jt}[\widehat{g}(\widehat{h}_{ljt-1}, r_{jt-1}) - \widehat{g}(\widehat{h}_{ljt-2}, r_{jt-2})]$, where the weights $\mu_{jt} = Y_{jt-2}/\sum_j Y_{jt-2}$ are given by the share of output of a firm two periods ago and we assume $E[\mu_{jt}\xi_{jt-1}|\omega_{jt-2}, r_{jt-2}] = 0$.[41] Columns (1)–(3) of Table 8 report this average for the entire sample and for the subsamples of observations with and without R&D.

As can be seen in Column (1) of Table 8, productivity growth is highest in some of the industries with high innovative activity (above 1.5% in industries 3 and 4 and around 2.5% in industry 6) followed by some of the industries with intermediate innovative activity (above 1.5% in industry 1).

Productivity growth is higher for performers than for non-performers in five industries, sometimes considerably so (Columns (2) and (3)). Taken together these industries account for over half of manufacturing output (see Appendix A for details). The decomposition into the contributions of observations with and without R&D to productivity growth shows that firms that perform R&D contribute between 65% and 85% of productivity growth in industries 3, 4, and 6 with high innovative activity and between 70% and 90% in industries 1 and 2 with intermediate innovative activity. This is all the more remarkable since in these industries between 20% and 45% of firms engage in R&D. While these firms manufacture between 45% and 75% of output, their contribution to productivity growth exceeds their share of output by an average 15%. That is, firms that engage in R&D tend not only to be larger than those that do not but also to grow even larger over time. Investments in R&D and related activities are thus a primary source of productivity growth.

6.4.   *Rate of return*

We finally compute an alternative—and perhaps more intuitive—measure of the return to R&D. The growth in expected productivity in Equation (21) can be decomposed as

$$\beta_t + g(\omega_{jt-1}, r_{jt-1}) - \omega_{jt-1} = [\beta_t + g(\omega_{jt-1}, r_{jt-1}) - g(\omega_{jt-1}, \underline{r})] + [g(\omega_{jt-1}, \underline{r}) - \omega_{jt-1}], \tag{22}$$

where $\underline{r}$ denotes a negligible amount of R&D expenditures.[42] The first term in brackets reflects the change in expected productivity that is attributable to R&D expenditures r_{jt-1}, the second

---

41. This assumption is plausible because the value of ξ_{jt-1} is not known to the firm when it makes the decisions that determine Y_{jt-2} and thus μ_{jt}.

42. Recall from Equation (13) that we allow the conditional expectation function $g(\cdot)$ to be different when the firm adopts the corner solution of zero R&D expenditures and when it chooses positive R&D expenditures. To avoid this discontinuity, we take $g(\omega_{jt-1}, \underline{r})$ to be a weighted average of $g_0(\omega_{jt-1})$ and $g_1(\omega_{jt-1}, \underline{r})$, where $\underline{r}$ is the 2.5th percentile of the industry's R&D expenditures.

TABLE 8

Productivity growth, rates of return to R&D and investment in physical capital, and degree of uncertainty

Industry	Productivity growth			Rates of return					Regr. of $\ln\left(\dfrac{\xi_{jt}^2}{Var(\omega_{jt})}\right)$ on[b]		Knowledge cap. mod. Net rate (s. e.)[b]
				R&D[a]			Capital Net rate	Ratio			
	Total	R&D (%)	No R&D (%)	Gross rate	Net rate	Depr.			R&D (s. e.)	Inv. (s. e.)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. Metals and metal products	0.0172	0.0225 (73.0)	0.0099 (27.0)	1.092	0.659	0.432	0.193	3.4	0.392 (0.164)	−0.326 (0.202)	2.025 (1.435)
2. Non-metallic minerals	0.0048	0.0101 (90.5)	0.0018 (9.5)	0.919	0.632	0.287	0.311	2.0	0.673 (0.221)	−0.018 (0.279)	0.763 (0.447)
3. Chemical products	0.0172	0.0210 (83.0)	0.0095 (17.0)	0.507	0.347	0.160	0.199	1.7	0.612 (0.141)	0.014 (0.258)	0.540 (0.249)
4. Agric. and ind. machinery	0.0157	0.0137 (63.1)	0.0234 (36.9)	0.745	0.296	0.449	0.247	1.2	0.772 (0.197)	0.123 (0.258)	0.685 (0.051)
6. Transport equipment	0.0266	0.0281 (75.7)	0.0237 (24.3)	0.848	0.480	0.368	0.168	2.9	0.736 (0.207)	−0.225 (0.364)	0.558 (0.211)
7. Food, drink, and tobacco	0.0019	−0.0033	0.0054	0.584	0.098	0.485	0.072	1.4	1.237 (0.164)	0.272 (0.206)	0.705 (0.184)
8. Textile, leather, and shoes	0.0108	0.0049 (12.4)	0.0146 (87.6)	0.918	0.332	0.585	0.075	4.4	1.587 (0.172)	0.189 (0.184)	0.681 (0.579)
9. Timber and furniture	0.0102	0.0386 (46.4)	0.0062 (53.6)	0.788	0.384	0.405	0.279	1.4	2.072 (0.382)	−0.236 (0.212)	0.093 (0.628)
10. Paper and printing products	0.0089	0.0015 (11.4)	0.0110 (88.6)	1.049	0.444	0.605	0.158	2.8	1.284 (0.274)	−0.455 (0.233)	1.008 (0.098)

[a] We calculate the first two terms of the decomposition in Equation (22) and infer the third term. We trim 5% of observations at each tail of the distribution of h_{ijt-1}. We further trim up to 30% of observations according to their rates.

[b] All standard errors are robust to heteroskedasticity and autocorrelation.

the change that takes place in the absence of investment in knowledge. That is, the second term in brackets is attributable to depreciation of already attained productivity and, consequently, is expected to be negative. The net effect of R&D is thus the sum of its gross effect (first term in brackets) and the impact of depreciation (second term).

Consider the change in expected productivity that is attributable to R&D expenditures. Multiplying $\beta_t + g(\omega_{jt-1}, r_{jt-1}) - g(\omega_{jt-1}, \underline{r})$ in Equation (22) by a measure of expected value added, say V_{jt}, gives the rent that the firm can expect from this investment at the time it makes its decisions. Dividing it further by R&D expenditures R_{jt-1} gives an estimate of the gross rate of return, or dollars obtained by spending one dollar on R&D.[43] Note that we compute the gross rate of return on R&D using value added instead of gross output both to make it comparable to the existing literature (*e.g.* Nadiri, 1993; Griliches and Regev, 1995; Griliches, 2000) and because value added is closer to profits than gross output. We similarly compute the net rate of return to R&D and the compensation for depreciation from the growth decomposition in Equation (22) by multiplying and dividing through by V_{jt} and R_{jt-1}.

Columns (4)–(6) of Table 8 summarize the gross rate of return to R&D and its decomposition into the net rate and the compensation for depreciation. We report weighted averages where the weights $\mu_{jt} = R_{jt-2} / \sum_j R_{jt-2}$ are given by the share of R&D expenditures of a firm two periods ago. The gross rate of return to R&D far exceeds the net rate, thus indicating that a large part of firms' R&D expenditures is devoted to maintaining already attained productivity rather than to advancing it. The net rates of return to R&D are around 40% and differ across industries, ranging from very modest values near 10% to 65%.[44] Interestingly enough, the net rate of return to R&D is higher in an industry where a large part of the variance in productivity is due to random shocks, as can be seen in Figure 4. This suggests that the net rate of return to R&D includes a compensation for the uncertainties inherent in the R&D process.

As a point of comparison we report the net rate of return on investment in physical capital in Column (7) of Table 8. We first compute the gross rate of return as $\beta_k V_{jt}/K_{jt}$ and then subtract the rate of depreciation of physical capital from it to obtain the net rate of return.[45] Column (8) reports the ratio of the net rates of return to R&D and investment in physical capital. Returns to R&D are clearly higher than returns to investment in physical capital. The net rate of return to R&D is often twice that of the net rate of return to investment in physical capital.

Recall that in our model the productivity innovation ξ_{jt} may be thought of as the realization of the uncertainties that are naturally linked to productivity plus the uncertainties inherent in the R&D process such as chance in discovery and success in implementation. The question therefore is whether an investment in knowledge indeed injects further uncertainties into the productivity process that would be absent if the firm did not engage in R&D. As before we measure the degree of uncertainty by the ratio of the variance of the productivity innovation ξ_{jt} to the variance of actual productivity ω_{jt}. Regressing an estimate of the log of the ratio

43. The average rate that we compute is close to the marginal rate of return to R&D. To see this, linearly approximate $g(\omega_{jt-1}, \ln \underline{R}) \simeq g(\omega_{jt-1}, \ln R_{jt-1}) + \frac{\partial g(\omega_{jt-1}, \ln R_{jt-1})}{\partial r_{jt-1}} \frac{1}{R_{jt-1}}(\underline{R} - R_{jt-1})$. If $\underline{R} \to 0$, then $g(\omega_{jt-1}, r_{jt-1}) - g(\omega_{jt-1}, \underline{r}) \equiv g(\omega_{jt-1}, \ln R_{jt-1}) - g(\omega_{jt-1}, \ln \underline{R}) \simeq \frac{\partial g(\omega_{jt-1}, r_{jt-1})}{\partial r_{jt-1}}$. In practice, we use firm-specific averages of value added and investment in knowledge.

44. As pointed out by an anonymous referee, reported R&D expenditures may be the tip of the iceberg in terms of the resources a firm devotes to maintaining and advancing its productivity. Because we divide the various terms in Equation (22) by R_{jt-1}, rates of return may be inflated by underreporting.

45. The rate of depreciation that is assumed in computing the stock of physical capital is around 0.1 but differs across industries and groups of firms within industries. We report a weighted average where the weights $\mu_{jt} = V_{jt} / \sum_j V_{jt}$ are given by are given by the share of value added of a firm. In practice, we use firm-specific averages of value added and the stock of physical capital.
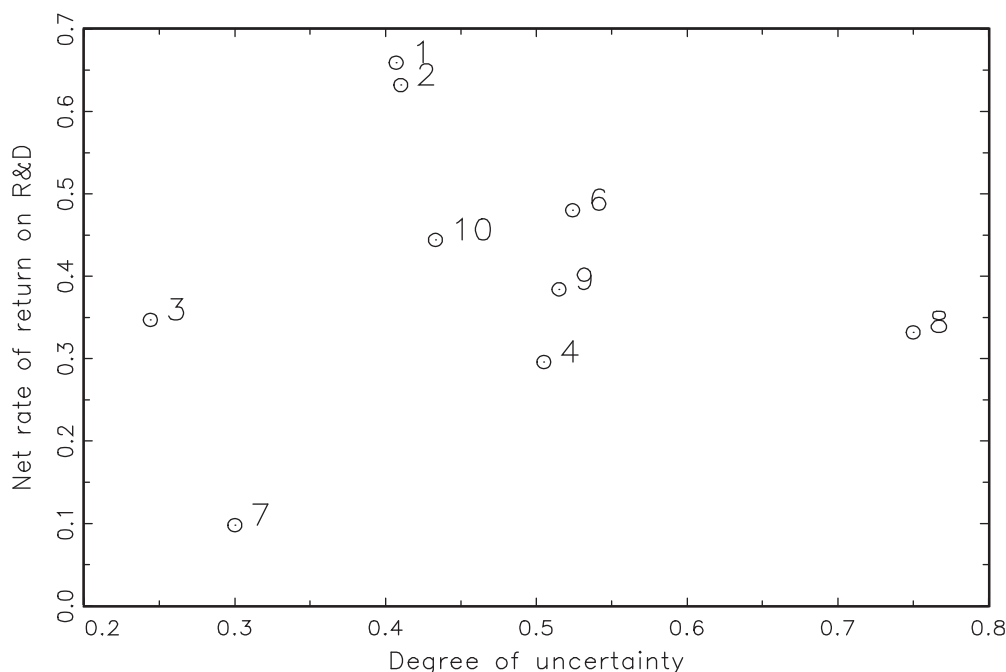
FIGURE 4
Return to R&D and uncertainty

$\xi_{jt}^2/Var(\omega_{jt})$ on a constant, a dummy for large firms with more than 200 workers, a dummy for investment in knowledge, and a dummy for investment in physical capital shows a positive impact of investment in knowledge on the degree of uncertainty in all industries (see Column (9) of Table 8).[46] Whereas investment in knowledge substantially increases the degree of uncertainty in the productivity process, investment in physical capital leaves it unchanged (Column (10)).

In sum, investment in knowledge is systematically more uncertain than investment in physical capital. The net rate of return includes a compensation for the uncertainties inherent in the R&D process. Moreover, the large gap between the net rates of return to R&D and investment in physical capital suggests that the uncertainties inherent in the R&D process are economically significant and matter for firms' investment decisions.

Comparison to knowledge capital model: To facilitate the comparison with the existing literature, we have used the value-added version of the knowledge capital model in Equation (20) to estimate the rate of return to R&D by regressing the first-difference of the log of value added on the first-differences of the logs of conventional inputs and the ratio R_{jt-1}/V_{jt-1} of R&D expenditures to value added.[47] The estimated coefficient of this ratio can be interpreted as the gross rate of return to R&D.[48] We obtain the net rate of return to R&D by subtracting the rate

46. We estimate $Var(\omega_{jt})$ separately for firms that do not engage in R&D, firms that engage in R&D and have R&D expenditures below the median and those that have R&D expenditures above the median.

47. To improve the estimates we impose the widely accepted constraint of constant returns to scale in the conventional inputs.

48. Recall that ε is the elasticity of value added with respect to knowledge capital. Since $\varepsilon\Delta c_{jt} = \frac{\partial V}{\partial C}\frac{C_{jt-1}}{V_{jt-1}}\Delta c_{jt} \simeq \frac{\partial V}{\partial C}\frac{\Delta C_{jt}}{V_{jt-1}}$ and R_{jt-1} approximates ΔC_{jt} (by the law of motion for knowledge capital), the estimated coefficient is $\frac{\partial V}{\partial C}$. Since

of depreciation of knowledge capital. As can be seen from Column (11) of Table 8, the net rates in the knowledge capital model are around 80%. While they are imprecisely estimated, they are higher than the net rates in our model with the exception of industry 9.

The knowledge capital literature has had limited success in estimating the rate of return to R&D. Griliches (2000) contends that "[e]arly studies of this topic were happy to get the sign of the R&D variable 'right' and to show that it matters, that it is a 'significant' variable, contributing to productivity growth" (p. 51). Estimates of the rate of return to R&D tended to be high, often implausibly high: "our current quantitative understanding of this whole process remains seriously flawed ... [T]he size of the effects we have estimated may be seriously off, perhaps by an order of magnitude" (Griliches, 1995, p. 83). Our estimates, by contrast, are more modest.

7. CONCLUDING REMARKS

In this article, we develop a model of endogenous productivity change resulting from investment in knowledge. While the knowledge capital model in its basic form can be viewed as a special case of our model, we differ in that, rather than attempting to construct a stock of knowledge capital from a firm's observed R&D expenditures, we consider productivity to be unobservable to the econometrician. We also derive an estimator for production functions in this setting.

Applying our approach to an unbalanced panel of more than 1800 Spanish manufacturing firms in nine industries during the 1990s, we show that the link between R&D and productivity is subject to a high degree of uncertainty, nonlinearity, and heterogeneity. R&D is a major determinant of the differences in productivity across firms and the evolution of firm-level productivity over time.

While our focus is on the link between R&D and productivity, we hope that our model of endogenous productivity change facilitates further inquiries into the determinants of productivity. To date, it has been applied by Aw *et al.* (2011) to examine the impact of export market participation on the productivity of firms; other applications include Maican and Orth (2008) and Añon and Manjon (2009).

Our parametric inversion and the resulting estimator for production functions may also prove useful in empirical research that requires multi-dimensional productivity measures. It is often possible to recover multiple unobservables from observing multiple decisions of a firm. Doraszelski and Jaumandreu (2009), for example, separately measure Hicks-neutral and labour-augmenting productivity to inquire about the nature of technological change. Another promising avenue for future research may be to study the impact of investments in R&D and related activities on product innovations. To this end, one may build on Aw *et al.* (2011) and Jaumandreu and Mairesse (2010) by introducing an additional unobservable that captures shifts and rotations in demand resulting from product innovations. In parallel to the productivity measure in the current article, this demand-side unobservable can then be endogenized by letting its law of motion depend on R&D expenditures.

APPENDIX A

We observe firms for a maximum of ten years between 1990 and 1999. We restrict the sample to firms with at least two years of data on all variables required for estimation. Because of data problems we exclude industry 5 (office and data-processing machines and electrical goods). Our final sample covers 1870 firms in 9 industries. The number of firms with 2, 3,…, 10 years of data is 259, 377, 244, 191, 171, 134, 127, 153, and 214, respectively. Table A1 gives the industry definitions along with their equivalent definitions in terms of the ESEE, National Accounts, and ISIC classifications

spending one dollar on R&D adds one unit of knowledge capital $\frac{\partial V}{\partial C}$ is, in turn, equal to $\frac{\partial V}{\partial R}$ or the gross rate of return to R&D.

TABLE A1

Industry definitions, equivalent classifications, and shares

Industry	Classifications			Share of value added	Number of subindustries
	ESEE	National accounts	ISIC		
	(1)	(2)	(3)	(4)	(5)
1. Ferrous and non-ferrous metals and metal products	1+4	DJ	D 27+28	12.6	13
2. Non-metallic minerals	2	DI	D 26	7.8	8
3. Chemical products	3+17	DG-DH	D 24+25	13.7	8
4. Agricultural and industrial machinery	5	DK	D 29	5.9	8
6. Transport equipment	8+9	DM	D 34+35	11.0	10
7. Food, drink, and tobacco	10+11+12	DA	D 15+16	16.5	10
8. Textile, leather, and shoes	13+14	DB-DC	D 17+18+19	7.9	13
9. Timber and furniture	15	DD-DN 38	D 20+30	6.3	8
10. Paper and printing products	16	DE	D 21+22	8.2	5
Total				90.0	

(Columns (1)–(3)). We finally report the shares of the various industries in the total value added of the manufacturing sector in 1995 (Column (4)).

The ESEE survey provides information on the total R&D expenditures of each firm in each year. These include the cost of intramural R&D activities, payments for outside R&D contracts with laboratories and research centers, and payments for imported technology in the form of patent licensing or technical assistance, with the various expenditures defined according to the OECD Oslo and Frascati manuals. We consider a firm to be performing R&D if it reports positive expenditures. While total R&D expenditures vary widely across firms, it is quite likely even for small firms that they exceed non-negligible values relative to firm size. In addition, firms are asked to provide many details about the combination of R&D activities, R&D employment, R&D subsidies, and the number of process and product innovations as well as the patents that result from these activities. Taken together, this supports the notion that the reported expenditures are truly R&D related.

In what follows we define the remaining variables.

- *Investment.* Value of current investments in equipment goods (excluding buildings, land, and financial assets) deflated by the price index of investment. By measuring investment in operative capital we avoid some of the more severe measurement issues of the other assets. We follow Ackerberg *et al.* (2007) and assume that the investment decided in period $t-1$ coincides with the investment observed in period t. Experimentation with the lagged value of this flow gave very similar results.
- *Capital.* Capital at current replacement values $\widetilde{K}_{jt}$ is computed recursively from an initial estimate and the data on current investments in equipment goods $\widetilde{I}_{jt}$. We update the value of the past stock of capital by means of the price index of investment P_{It} as $\widetilde{K}_{jt} = (1-\delta)\frac{P_{It}}{P_{It-1}}\widetilde{K}_{jt-1} + \widetilde{I}_{jt-1}$, where δ is an industry-specific estimate of the rate of depreciation. Capital in real terms is obtained by deflating capital at current replacement values by the price index of investment as $K_{jt} = \frac{\widetilde{K}_{jt}}{P_{It}}$.
- *Labour.* Total hours worked computed as the number of workers times the average hours per worker, where the latter is computed as normal hours plus average overtime minus average time lost at the workplace.
- *Materials.* Value of intermediate consumption (including raw materials, components, energy, and services) deflated by a firm-specific price index of materials.
- *Output.* Value of produced goods and services computed as sales plus the variation of inventories deflated by a firm-specific price index of output.
- *Price of investment.* Equipment goods component of the index of industry prices computed and published by the Spanish Ministry of Industry.
- *Wage.* Hourly wage cost computed as total labour cost including social security payments divided by total hours worked.
- *Price of materials.* Firm-specific price index for intermediate consumption. Firms are asked about the price changes that occurred during the year for raw materials, components, energy, and services. The price index is computed as a Paasche-type index of the responses and normalized by the average of its values for each firm.

- *Price of output.* Firm-specific price index for output. Firms are asked about the price changes they made during the year in up to 5 separate markets in which they operate. The price index is computed as a Paasche-type index of the responses and normalized by the average of its values for each firm.
- *Index of market dynamism.* Firms are asked to assess the current and future situation (contraction, stability, or expansion) of up to 5 separate markets in which they operate. The index of market dynamism is computed as a weighted average of the responses and proxies for the demand shifter d_{jt}.
- *Technological sophistication.* Dummy variable that takes the value one if the firm uses digitally controlled machines, robots, CAD/CAM, or some combination of these procedures.
- *Identification between ownership and control.* Dummy variable that takes the value one if the owner of the firm or the family of the owner hold management positions.
- *Age.* Years elapsed since the foundation of the firm with a maximum of 40 years.
- *Work force.* Fraction of workers under fixed term contracts with very small or no severance pay (temporary workers). Fraction of non-production employees (white collar workers), workers with an engineering degree (engineers), and workers with an intermediate degree (technicians).
- *Location of industrial employment.* Nineteen dummy variables corresponding to the Spanish autonomous communities when employment is located in a unique region and another dummy variable when employment is spread over several regions.
- *Product submarket.* Dummy variables corresponding to a finer breakdown of the 9 industries into subindustries (see Column (5) of Table A1).

APPENDIX B

To simplify the exposition, we restrict the production function in Equation (1) to $y_{jt} = \beta_l l_{jt} + \omega_{jt} + e_{jt}$, the law of motion in Equation (3) to an exogenous $AR(1)$ process with $g(\omega_{jt-1}) = \rho \omega_{jt-1}$, and assume perfect competition in the product market with $\eta(\cdot) = \infty$. Normalizing the price of output, the labour demand function is

$$l_{jt} = \frac{1}{1-\beta_l}\left(\ln \beta_l + \mu - w_{jt}^* + \omega_{jt}\right). \tag{B.1}$$

While Equation (B.1) accurately describes the firm's decision-making process, there is a problem of observability. Assume, as may easily happen in practice, that we as econometricians imperfectly observe prices: instead of w_{jt}^* we observe w_{jt}, where the difference between observed and true prices is measurement error $v_{jt} = w_{jt} - w_{jt}^*$. Given what we observe, we can no longer recover true productivity as

$$h_{ljt}^* = \omega_{jt} = \lambda_l + (1-\beta_l)l_{jt} + w_{jt}^*. \tag{B.2}$$

where $\lambda_l = -\ln \beta_l - \mu$. Instead we recover

$$h_{ljt} = \lambda_l + (1-\beta_l)l_{jt} + w_{jt} = h_{ljt}^* + v_{jt}. \tag{B.3}$$

We refer to h_{ljt} as observed (as opposed to true) productivity and base the estimation on it.

We assume throughout that the measurement error v_{jt} is independent of productivity ω_{jt} and distinguish two polar cases: in the case of *classical* measurement error with respect to prices, the measurement error v_{jt} is uncorrelated with true prices w_{jt}^* and therefore correlated with observed prices w_{jt}; in the case of *non-classical* measurement error with respect to prices, the measurement error v_{jt} is uncorrelated with observed prices w_{jt} and therefore correlated with true prices w_{jt}^*.

To explore the consequences of imperfectly observable prices, we assume that

$$v_{jt} = E\left[v_{jt}|l_{jt}, w_{jt}\right] + u_{jt} = \delta_0 + \delta_1 l_{jt} + \delta_2 w_{jt} + u_{jt}, \tag{B.4}$$

where u_{jt} is mean independent of l_{jt} and w_{jt}.[49] For the case of classical measurement error with respect to prices we then have

$$\delta_1 = (1-\beta_l)\frac{\sigma_v^2}{\sigma_{w^*}^2 \sigma_\omega^2 + \sigma_{w^*}^2 \sigma_v^2 + \sigma_\omega^2 \sigma_v^2}\sigma_{w^*}^2, \quad \delta_2 = \frac{\sigma_v^2}{\sigma_{w^*}^2 \sigma_\omega^2 + \sigma_{w^*}^2 \sigma_v^2 + \sigma_\omega^2 \sigma_v^2}(\sigma_{w^*}^2 + \sigma_\omega^2)$$

and

$$\delta_1 = (1-\beta_l)\frac{\sigma_v^2}{\sigma_v^2 + \sigma_\omega^2}, \qquad \delta_2 = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\omega^2}$$

for the case of non-classical measurement error with respect to prices.

49. Assumption (B.4) does not require v_{jt}, l_{jt}, and w_{jt} to be jointly normal distributed and extends, for example, to the elliptical and Pearson families. Moreover, it is easy to show that if the joint distribution of v_{jt}, ω_{jt}, and w_{jt}^* (classical measurement error) or w_{jt} (non-classical measurement error) is in the elliptical family, then so is the induced joint distribution of v_{jt}, l_{jt}, and w_{jt}. More generally, assumption (B.4) may be viewed as a useful approximation that allows us to explore the consequences of imperfectly observable prices.

Substituting $\omega_{jt} = \rho\omega_{jt-1} + \xi_{jt}$ and Equation (B.2) into $y_{jt} = \beta_l l_{jt} + \omega_{jt} + e_{jt}$, the model is

$$y_{jt} = \lambda_0 + \beta_l l_{jt} + \rho h^*_{ljt-1} + \xi_{jt} + e_{jt},$$

where $\lambda_0 = -\rho(\ln\beta_l + \mu)$. Using Equations (B.3) and (B.4) to express the model in terms of observables we thus have

$$y_{jt} = \lambda_0 + \beta_l l_{jt} + \rho\left(h_{ljt-1} - \delta_0 - \delta_1 l_{jt-1} - \delta_2 w_{jt-1}\right) - \rho u_{jt-1} + \xi_{jt} + e_{jt}$$

$$= -\tilde{\lambda}_0 + \beta_l l_{jt} + \tilde{\rho}\left((1 - \tilde{\beta}_l)l_{jt-1} + w_{jt-1}\right) - \rho u_{jt-1} + \xi_{jt} + e_{jt},$$

where

$$\tilde{\lambda}_0 = \lambda_0 - \rho\delta_0, \quad \tilde{\rho} = \rho(1 - \delta_2), \quad 1 - \tilde{\beta}_l = \frac{1 - \beta_l - \delta_1}{1 - \delta_2}.$$

We conclude that imperfectly observable prices drive a wedge between the coefficients in the production function and their counterparts in the inverse labour demand function which may cause the parameter restrictions test to reject.[50]

Fortunately, however, by simply dropping the restriction between parameters it remains possible to consistently estimate the coefficients in the production function (β_l in this simple example) because u_{jt-1} is uncorrelated with the instruments (constant, l_{jt-1}, and w_{jt-1}). With these coefficients in hand we are moreover able to consistently recover actual productivity ω_{jt}. What we are no longer able to do is consistently decompose ω_{jt} into expected productivity $\rho\omega_{jt-1}$ and the random shock ξ_{jt}.[51]

APPENDIX C

Introducing adjustment costs $C_{LP}(L_{Pjt}, L_{Pjt-1})$ the Bellman equation (2) becomes

$$V(s_{jt}) = \max_{i_{jt}, r_{jt}, L_{Pjt}} \Pi(s_{jt}) - C_i(i_{jt}) - C_r(r_{jt}) - C_{LP}(L_{Pjt}, L_{Pjt-1}) + \frac{1}{1+\rho}E\left[V(s_{jt+1})|s_{jt}, i_{jt}, r_{jt}\right],$$

where the indirect profit function $\Pi(\cdot)$ is based on the production function in Equation (15) and the vector of state variables $s_{jt} = (t, k_{jt}, L_{Pjt-1}, \omega_{jt}, w_{Pjt}, w_{Tjt}, p_{Mjt}, d_{jt})$ includes lagged permanent labour L_{Pjt-1}, the log of the wage of permanent workers w_{Pjt}, and that of temporary workers w_{Tjt}.

At an interior solution the first-order condition for permanent labour is

$$\beta_l\alpha e^{\beta_0}e^{\beta_t t}K_{jt}^{\beta_k}L_{Pjt}^{\beta_l\alpha-1}L_{Tjt}^{\beta_l(1-\alpha)}M_{jt}^{\beta_m}e^{\omega_{jt}}e^\mu = \frac{W_{Pjt}\left(1+\Delta_{jt}\right)}{P_{jt}\left(1-\frac{1}{\eta(p_{jt},d_{jt})}\right)}, \tag{C.1}$$

where, by the envelope theorem, the gap between the wage of permanent workers W_{Pjt} and their shadow wage is

$$\Delta_{ij} = \frac{1}{W_{Pjt}}\left(\frac{\partial C_{LP}(L_{Pjt}, L_{Pjt-1})}{\partial L_{Pjt}} - \frac{1}{1+\rho}E\left[\frac{\partial V(s_{jt+1})}{\partial L_{Pjt}}|s_{jt}, i_{jt}, r_{jt}\right]\right)$$

$$= \frac{1}{W_{Pjt}}\left(\frac{\partial C_{LP}(L_{Pjt}, L_{Pjt-1})}{\partial L_{Pjt}} + \frac{1}{1+\rho}E\left[\frac{\partial C_{LP}(L_{Pjt+1}, L_{Pjt})}{\partial L_{Pjt}}|s_{jt}, i_{jt}, r_{jt}\right]\right). \tag{C.2}$$

At an interior solution the first-order conditions for temporary labour and materials are

$$\beta_l(1-\alpha)e^{\beta_0}e^{\beta_t t}K_{jt}^{\beta_k}L_{Pjt}^{\beta_l\alpha}L_{Tjt}^{\beta_l(1-\alpha)-1}M_{jt}^{\beta_m}e^{\omega_{jt}}e^\mu = \frac{W_{Tjt}}{P_{jt}\left(1-\frac{1}{\eta(p_{jt},d_{jt})}\right)}, \tag{C.3}$$

$$\beta_m e^{\beta_0}e^{\beta_t t}K_{jt}^{\beta_k}L_{Pjt}^{\beta_l\alpha}L_{Tjt}^{\beta_l(1-\alpha)}M_{jt}^{\beta_m-1}e^{\omega_{jt}}e^\mu = \frac{P_{Mjt}}{P_{jt}\left(1-\frac{1}{\eta(p_{jt},d_{jt})}\right)}. \tag{C.4}$$

50. For the case of non-classical measurement error with respect to prices, we obtain $1 - \tilde{\beta}_l = \frac{1-\beta_l-\delta_1}{1-\delta_2} = 1 - \beta_l$ because $\delta_1 = (1-\beta_l)\delta_2$ in our simple example, but this does not generalize to the model with capital.

51. Our conclusions generalize beyond the simple example if we strengthen assumption (B.4). For example, for an exogenous Markov process for productivity with law of motion $\omega_{jt} = \rho_1\omega_{jt-1} + \rho_2\omega_{jt-1}^2 + \rho_3\omega_{jt-1}^3 + \xi_{jt}$ we further assume that u_{jt}^2 and u_{jt}^3 are mean independent of l_{jt} and w_{jt}, i.e. that u_{jt} is homoskedastic and has constant skewness conditional of l_{jt} and w_{jt}.

Multiplying the first-order conditions (C.1) and (C.3) by L_{Pjt} and L_{Tjt}, respectively, adding them, and rearranging yields an expression resembling a first-order condition for effective labour $L_{jt}^* = L_{Pjt}^{\alpha} L_{Tjt}^{(1-\alpha)}$:

$$\beta_l e^{\beta_0} e^{\beta_t t} K_{jt}^{\beta_k} \left(L_{jt}^*\right)^{\beta_l-1} M_{jt}^{\beta_m} e^{\omega_{jt}} e^{\mu} = \frac{W_{Pjt}(1+\Delta_{ij})L_{Pjt}+W_{Tjt}L_{Tjt}}{P_{jt}\left(1-\frac{1}{\eta(p_{jt},d_{jt})}\right)L_{jt}^*}.$$

Rewriting in terms of observed labour $L_{jt} = L_{Pjt} + L_{Tjt}$ and the observed wage $W_{jt} = W_{Pjt}(1-S_{Tjt}) + W_{Tjt}S_{Tjt}$, where $S_{Tjt} = \frac{L_{Tjt}}{L_{jt}}$ is the share of temporary workers in our data, yields

$$\beta_l e^{\beta_0} e^{\beta_t t} K_{jt}^{\beta_k} (L_{jt}C_{1jt})^{\beta_l-1} M_{jt}^{\beta_m} e^{\omega_{jt}} e^{\mu} = \frac{W_{jt}}{P_{jt}\left(1-\frac{1}{\eta(p_{jt},d_{jt})}\right)} \frac{C_{2jt}}{C_{1jt}}, \tag{C.5}$$

where the corrections C_{1jt} and C_{2jt} are defined in Equations (17) and (18). Together with Equation (C.4), Equation (C.5) yields the aggregate labour demand function in Equation (16). We finally obtain Equation (19) by dividing the first-order conditions (C.1) and (C.3) and recalling that $S_{Tjt} = \frac{L_{Tjt}}{L_{jt}}$.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

AÑON, D. and MANJON, M. (2009), "Internationalization, R&D and Productivity" (Working paper, University of Valencia, Valencia).

ACKERBERG, D., BENKARD, L., BERRY, S. and PAKES, A. (2007), "Econometric Tools for Analyzing Market Outcomes", in Heckman, J. and Leamer, E. (eds), 'Handbook of Econometrics', Vol. 6A (Amsterdam: North-Holland) 4171–4276.

ACKERBERG, D., CAVES, K. and FRAZER, G. (2006), "Structural Identification of Production Functions" (Working paper, UCLA, Los Angeles).

ADDA, J. and COOPER, R. (2003), *Dynamic Economics: Quantitative Methods and Applications* (Cambridge: MIT Press).

AGUIRREGABIRIA, V. and ALONSO-BORREGO, C. (2009), "Labor Contracts and Flexibility: Evidence from a Labor Market Reform in Spain" (Working paper no. 3478, University of Toronto, Toronto).

AI, C. and CHEN, X. (2003), "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions", *Econometrica*, **71**, 1795–1843.

AI, C. and CHEN, X. (2007), "Estimation of Possibly Misspecified Semiparameteric Conditional Moment Restriction Models with Different Conditioning Variables", *Journal of Econometrics*, **141**, 5–43.

AMEMIYA, T. (1974), "The Nonlinear Two-stage Least-squares Estimator", *Journal of Economic Literature*, **2**, 105–110.

ANTRÀS, P. and HELPMAN, E. (2004), "Global Sourcing", *Journal of Political Economy*, **112**, 552–580.

ARELLANO, M. (2003), *Panel Data Econometrics* (Oxford: Oxford University Press).

AW, B., ROBERTS, M. and XU, D. (2011), "R&D Investment, Exporting, and Productivity Dynamics", *American Economic Review*, **101**, 1312—1344.

BARRETT, G. and DONALD, S. (2003), "Consistent Tests for Stochastic Dominance", *Econometrica*, **71**, 71–104.

BENEITO, P. (2001), "R&D, Productivity and Spillovers at the Firm Level: Evidence from Spanish Panel Data", *Investigaciones Economicas*, **25**, 289–313.

BERRY, S., LEVINSOHN, J. and PAKES, A. (1995), "Automobile Prices in Market Equilibrium", *Econometrica*, **63**, 841–890.

BLUNDELL, R. and BOND, S. (2000), "GMM Estimation with Persistent Panel Data: An Application to Production Functions", *Econometric Reviews*, **19**, 321–340.

BOND, S. and SÖDERBOM, M. (2005), "Adjustment Costs and the Identification of Cobb-Douglas Production Functions" (Working paper WP05/04, The Institute for Fiscal Studies, London).

BUETTNER, T. (2005), "R&D and the Dynamics of Productivity" (Working paper, LSE, London).

CHAMBERS, R. (1988), *Applied Production Analysis: A Dual Approach* (Cambridge: Cambridge University Press).

COOPER, R. and WILLIS, J. (2004), "A Comment on the Economics of Labor Adjustment: Mind the Gap", *American Economic Review*, **84**, 1223–1237.

DE LOECKER, J. (2010), "A Note on Detecting Learning by Exporting" (Working paper no. 16548, NBER, Cambridge).

DELGADO, M., FARIÑAS, J. and RUANO, S. (2002), "Firm Productivity and Export Markets: A non-Parametric approach", *Journal of International Economics*, **57**, 397–422.

DJANKOV, S., LA PORTA, R., LOPEZ-DE SILANES, F. and SHLEIFER, A. (2002), "The Regulation of Entry", *Quarterly Journal of Economics*, **117**, 1–37.

DOLADO, J., GARCIA-SERRANO, C. and JIMENO, J. (2002), "Drawing Lessons from the Boom of Temporary Jobs in Spain", *Economic Journal*, **112**, 270–295.

DORASZELSKI, U. and JAUMANDREU, J. (2009), "Measuring the Bias of Technological Change" (Working paper, Harvard University, Cambridge).

ESLAVA, M., HALTIWANGER, J., KUGLER, A. and KUGLER, M. (2004), "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia", *Journal of Development Economics*, **75**, 333—371.

EUROPEAN COMMISSION (2001), "Statistics on Innovation in Europe" (Enterprise DG, Brussels).

EUROPEAN COMMISSION (2004), "European Competitiveness Report" (Enterprise DG, Brussels).

FOSTER, L., HALTIWANGER, J. and SYVERSON, C. (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?", *American Economic Review*, **98**, 394–425.

GONZALEZ, X., JAUMANDREU, J. and PAZO, C. (2005), "Barriers to Innovation and Subsidy Effectiveness", *Rand Journal of Economics*, **36**, 930–950.

GRILICHES, Z. (1979), "Issues in Assessing the Contribution of R&D to Productivity Growth", *Bell Journal of Economics*, **10**, 92–116.

GRILICHES, Z. (1995), "R&D and Productivity", in Stoneman, P. (ed.) *Handbook of the Economics of Innovation and Technical Change* (Oxford: Blackwell).

GRILICHES, Z. (1998), *R&D and Productivity: The Econometric Evidence* (Chicago; University of Chicago Press).

GRILICHES, Z. (2000), *R&D, Education, and Productivity: A Retrospective* (Cambridge: Harvard University Press).

GRILICHES, Z. and MAIRESSE, J. (1998), "Production Functions: The Search for Identification", in Strom, S. (ed.) *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium* (Cambridge: Cambridge University Press).

GRILICHES, Z. and REGEV, H. (1995), "Firm Productivity in Israeli industry 1979–1988", *Journal of Econometrics*, **65**, 175–203.

GROSSMAN, G. and ROSSI-HANSBERG, E. (2008), "Trading Tasks: A Simple Theory of Offshoring", *American Economic Review*, **95**, 1978–1997.

HALL, B. and HAYASHI, F. (1989), "Research and Development as an Investment" (Working paper no. 2973, NBER, Cambridge).

HALL, B. and MAIRESSE, J. (1995), "Exploring the Relationship between R&D and Productivity in French Manufacturing Firms", *Journal of Econometrics*, **65**, 263–293.

HALL, B., MAIRESSE, J. and MOHNEN, P. (2010), "Measuring the Returns to R&D", in Hall, B. and Rosenberg, N. (eds), *Handbook of the Economics of Innovation* (Amsterdam: North-Holland).

HANSEN, L. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, **50**, 1029–1054.

HELPMAN, E., MELITZ, M. and YEAPLE, S. (2004), "Export versus FDI with Heterogeneous Firms", *American Economic Review*, **94**, 300–316.

HUANG, G. and HU, Y. (2011), "Estimating Production Functions with Robustness against Errors in the Proxy Variables" (Working paper, Yale University, New Haven).

JAUMANDREU, J. and MAIRESSE, J. (2010), "Innovation and Welfare: Results from Joint Estimation of Production and Demand Functions" (Working paper no. 16221, NBER, Cambridge).

KLETTE, T. (1996), "R&D, Scope Economies, and Plant Performance", *Rand Journal of Economics*, **27**, 502–522.

LEVINSOHN, J. and PETRIN, A. (2003), "Estimating Production Functions using Inputs to Control for Unobservables", *Review of Economic Studies*, **70**, 317–341.

LINTON, O., MAASOUMI, E. and WHANG, Y. (2005), "Consistent Tests of Stochastic Dominance under General Sampling Schemes", *Review of Economic Studies*, **72**, 735–765.

MAICAN, F. and ORTH, M. (2008), "Productivity Dynamics and the Role of "big-box" Entrants in Retailing" (Working paper, University of Gothenburg, Göteborg).

MAIRESSE, J. and SASSENOU, M. (1991), "R&D and Productivity: A Survey of Econometric Studies at the Firm Level" (Working paper no. 3666, NBER, Cambridge).

MARSCHAK, J. and ANDREWS, W. (1944), "Random Simultaneous Equations and the Theory of Production", *Econometrica*, **12**, 143–205.

NADIRI, M. (1993), "Innovations and Technical Spillovers" (Working paper, New York University, New York).

NADIRI, M. and PRUCHA, I. (1996), "Estimation of Depreciation Rate of Physical and R&D Capital in the U.S. Total Manufacturing Sector", *Economic Inquiry*, **34**, 45–56.

NEWEY, W. (1990), "Efficient Instrumental Variables Estimation of Nonlinear Models", *Econometrica*, **58**, 809–837.

NEWEY, W. (1993), "Efficient Estimation of Models with Conditional Moment Restrictions", in Maddala, G., Rao, C. and Vinod, H. (eds), *Handbook of Statistics*, Vol. 11 (Amsterdam: Elsevier) 419–454.

NEWEY, W., POWELL, J. and VELLA, F. (1999), "Nonparametric Estimation of Triangular Simultaneous Equations Models", *Econometrica*, **67**, 565–603.

OLLEY, S. and PAKES, A. (1996), "The Dynamics of Productivity in the Telecommunications Industry", *Econometrica*, **64**, 1263–1297.

ORNAGHI, C. (2006), "Spillovers in Product and Process Innovation: Evidence from Manufacturing Firms", *International Journal of Industrial Organization*, **24**, 349–380.

PAKES, A. and GRILICHES, Z. (1984), "Estimating Distributed Lags in Short Panels with an Application to the Specification of Depreciation Patterns and Capital Stock Constructs", *Review of Economic Studies*, **51**, 243–262.

PAKES, A. and SCHANKERMAN, M. (1984), "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources", in Griliches, Z. (ed.) *R&D, Patents and Productivity* (Chicago: University of Chicago Press).

RIVERS, D. and VUONG, Q. (2002), "Model Selection Tests for Nonlinear Dynamic Models", *Econometrics Journal*, **5**, 1–39.

ROBERTS, M. and SUPINA, D. (1996), "Output Price, Markups, and Producer Size", *European Economic Review*, **40**, 909–921.

ROBINSON, P. (1988), "Root-n-consistent Semiparametric Regression", *Econometrica*, **56**, 931–954.

SCHANKERMAN, M. and PAKES, A. (1986), "Estimates of the Value of Patent Rights in European Countries during the post-1950 Period", *Economic Journal*, **96**, 1052–1076.

WOOLDRIDGE, J. (2009), "On Estimating Firm-level Production Functions using Proxy Variables to Control for Unobservables", *Economics Letters*, **104**, 112–114.

WOOLDRIDGE, J. (2010), *Econometric Analysis of Cross Section and Panel Data*, 2nd edn (Cambridge: MIT Press).