R&D and productivity: Estimating endogenous productivity*

Ulrich Doraszelski†
Harvard University and CEPR

Jordi Jaumandreu‡
Boston University, Universidad Carlos III, and CEPR

April 7, 2009

Abstract

We develop a model of endogenous productivity change that accounts for investment in knowledge. Our dynamic investment model extends the tradition of the knowledge capital model of Griliches (1979) that has remained a cornerstone of the productivity literature for more than 25 years. We relax the assumptions on the R&D process and examine the impact of the investment in knowledge on the productivity of firms. We also derive an estimator for production functions in this setting.

We illustrate our approach on an unbalanced panel of more than 1800 Spanish manufacturing firms in nine industries during the 1990s. Our findings indicate that the link between R&D and productivity is subject to a high degree of uncertainty, nonlinearity, and heterogeneity across firms. By accounting for uncertainty and nonlinearity, we extend the knowledge capital model. Capturing heterogeneity 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.

1 Introduction

Firms invest in R&D and related activities to develop and introduce process and product innovations. By enhancing productivity these investments in knowledge create long-lived assets for firms, similar to their investments in physical capital. Our goal in this paper is to

*Earlier versions of this paper were circulated as "R&D and productivity: The knowledge capital model revisited" and "R&D and productivity: Estimating production functions when productivity is endogenous." We thank Dan Ackerberg, Michaela Draganska, David Greenstreet, Wes Hartmann, Dale Jorgenson, Ken Judd, Saul Lach, Jacques Mairesse, Harikesh Nair, Sridhar Narayanan, Ariel Pakes, Mark Roberts, Marc Rysman, Mark Schankerman, and Jim Tybout for helpful discussions and Laia Castany for excellent research assistance. Doraszelski gratefully acknowledges the hospitality of the Hoover Institution during the academic year 2006/07.

†Department of Economics, Harvard University, Littauer Center, 1875 Cambridge Street, Cambridge, MA 02138, USA. E-mail: doraszelski@harvard.edu.

‡Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, USA. E-mail: jordij@bu.edu.

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 that accounts for investment in knowledge. We also derive an estimator for production functions in this setting. With these tools in hand we study the relationship between R&D and productivity in Spanish manufacturing firms during the 1990s. We particularly pay attention to the uncertainties and nonlinearities in the R&D process and their implications for heterogeneity across firms.

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 capital stock of the firm as do the subsequent decisions on static inputs such as labor 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, and gains and losses due to changes in labor composition and management abilities. For firms that engage 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.

It has, of course, long been recognized that the productivity process is endogenous. Griliches (1979) proposed to augment the production function with the stock of knowledge as proxied for by a firm's past R&D expenditures. 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 and also extended to macroeconomic growth models. While useful as a practical tool, the knowledge capital model has a long list of known drawbacks (see Griliches 2000). The basic model assumes linear and certain accumulation of knowledge from period to period in proportion to R&D expenditures as well as linear and certain depreciation.[1]

Our dynamic investment model can be viewed as a generalization of the knowledge capital model in its basic—albeit most widely used—form or as a practical alternative to more general forms of the knowledge capital model. In particular, 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 that current and past

[1]There have been few attempts to relax these assumptions. Pakes & Schankerman (1984*a*) model the creation of knowledge by specifying a production function in terms of R&D capital and R&D labor. Pakes & Schankerman (1984*b*) and Nadiri & Prucha (1996) attempt to estimate the pattern of depreciation over time. Hall & Hayashi (1989) and Klette (1996) allow for nonlinearity in knowledge accumulation but retain stringent functional form assumptions. Hall & Hayashi (1989) also allow for uncertainty. Finally, Jaffe (1986) initiates ways of accounting for the appropriability of the external flows of knowledge or spillovers (for a recent example see Griffith, Harrison & Van Reenen 2006).

investments in knowledge are likely to interact with each other in many ways. Since there is little reason to believe that features such as complementarities in the accumulation of knowledge and economies of scale can be adequately captured by simple functional forms, we model the interactions between current and past investments in knowledge in a flexible fashion. Finally, we relax the assumption that the obsolescence of previously acquired knowledge can be described by a constant rate of depreciation. This allows us to more closely assess the impact of the investment in knowledge on the productivity of firms.

To retrieve productivity at the level of the firm, we have to estimate the parameters of the production function. A major obstacle is that the decisions that a firm makes depend on its productivity. Because the productivity of the firm is unobserved by the econometrician, this gives rise to an endogeneity problem (Marschak & Andrews 1944). Intuitively, if a firm adjusts to a change in its productivity by expanding or contracting its production depending on whether the change is favorable or not, then unobserved productivity and input usage are correlated and biased estimates result.[2]

Recent advances in the structural estimation of production functions, starting with the dynamic investment model of Olley & Pakes (1996) (OP), tackle this issue. 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. In addition to OP, this line of research includes contributions by Levinsohn & Petrin (2003) (LP) and Ackerberg, Caves & Frazer (2006) (ACF) as well as a long list of applications.

Common to OP, LP, and ACF is the assumption that any changes in its productivity are exogenous to the firm. But if productivity evolves independently of R&D, then a firm has no incentive to invest in R&D in the first place. This makes it difficult to study the link between R&D and productivity. Indeed, the available estimators have mostly been applied to analyze changes in productivity in response to exogenous shocks such as deregulation (e.g., OP) or trade liberalization (e.g., Pavcnik 2002, Topalova 2004).

Incorporating R&D expenditures into the dynamic investment model of OP in order to endogenize the productivity process is difficult because it requires severely restricting how R&D can impact productivity in order to guarantee that investment in physical capital can be inverted to back out productivity (Buettner (2005), see Section 3 for details). Instead of relying on the firm's dynamic programming problem, 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 resulting input demands are invertible functions of unobserved productivity. This enables us to control for productivity and obtain consistent estimates of the parameters of the production function. We build on the previous literature by recognizing that, given a parametric specification of the production function, the functional form of these inverse input demand functions is known. Because we fully exploit the structural as-

[2]See Griliches & Mairesse (1998) and Ackerberg, Benkard, Berry & Pakes (2007) for reviews of this and other problems that arise in the estimation of production functions.

sumptions, we do not have to rely on nonparametric methods to estimate the inverse input demand function. This yields a particularly simple estimator for production functions.

We apply our estimator to an unbalanced panel of more than 1800 Spanish manufacturing firms in nine industries during the 1990s. This broad coverage of industries is unusual, and it allows us to examine the link between R&D and productivity in a variety of settings that potentially differ in the importance of R&D.

The data refute the assumptions at the heart of the knowledge capital model. To begin with, the impact of current R&D on future productivity depends crucially on current productivity, and there is evidence of complementarities as well as increasing returns to R&D. This casts doubt on the linearity assumption in the accumulation and depreciation of knowledge. Furthermore, the R&D process must be treated as inherently uncertain. We estimate that, depending on the industry, between 20% and 60% of the variance in actual productivity is explained by productivity innovations that cannot be predicted when decisions on R&D expenditures are made. Our estimates imply that the return to R&D is often twice that of the return to investment in physical capital. This suggests 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. Whereas firms with the same time path of R&D expenditures have necessarily the same productivity in the knowledge capital model, in our setting this is no longer the case because we allow the shocks to productivity to accumulate over time. 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, the expected productivity of firms that perform R&D is systematically more favorable 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 takes a sort of average over firms with distinct innovative activities and hence blurs remarkable differences in the impact of the investment in knowledge on the productivity of firms. In addition, we estimate that the contribution of firms that perform R&D explains between 65% and 85% of productivity growth in the industries with intermediate or high innovative activity. R&D expenditures are thus a primary source of productivity growth.

Our analysis further implies that productivity is considerably more fluid than what the knowledge capital literature suggests. 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. Hidden behind these averages, however, is a substantial amount of heterogeneity across firms.

Overall, the link between R&D and productivity is subject to a high degree of uncertainty, nonlinearity, and heterogeneity across firms. Abstracting from uncertainty and nonlinearity, as is done in the knowledge capital model, or assuming an exogenous process for productivity, as is done in the literature following OP, overlooks some of its most interesting features.

R&D plays a key role in the debate about growth in Spain. A recent OECD study voiced concern that boosting R&D and innovation is a challenge for Spain given its industrial structure with few high-tech industries and mostly small and medium-sized firms (OECD 2007) and in 2005 the Spanish government launched the Ingenio 2010 initiative that is targeted at funding large-sized, high-risk research projects. Our findings have implications for industry dynamics and R&D policy that directly speak to these issues. While a fuller exploration is left to future research, we note here some tentative conclusions. First, our results imply that the scope for industrial change in Spain is limited. This is because firms are repeatedly subjected to shocks that make it hard for them to "break away" from their rivals and remain at or near the top of the productivity distribution. Second, because it captures the heterogeneity across firms, our model can be used to determine the allocation of subsidies, a major issue in R&D policy. Our results indicate a systematic relationship between firm size and the return to R&D, thereby suggesting that, if the goal is to maximize returns, then larger firms should be subsidized more extensively than smaller firms. At the same time, our results attest to the high degree of uncertainty in the link between R&D and productivity. But if uncertainty inhibits firms' investments in R&D, then a case can be made for R&D policy to be redirected from subsidizing the cost of R&D to providing insurance against particularly unfavorable outcomes.

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. The investment decisions are made in a discrete time setting with the goal of maximizing the expected net present value of future cash flows. The firm has the Cobb-Douglas production function

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

where y_{jt} is the log of output of firm j in period t, l_{jt} the log of labor, and k_{jt} the log of capital. We follow the convention that lower case letters denote logs and upper case letters levels and focus on a value-added specification to simplify the exposition. Capital is the only dynamic input among the conventional factors of production, and accumulates according to $K_{jt} = (1-\delta)K_{jt-1} + I_{jt-1}$. 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}.

We follow OP and often refer to ω_{jt} as "unobserved productivity" since it is unobserved from the point of view of the econometrician (but known to the firm). Productivity is presumably highly correlated over time and perhaps also across firms. In contrast, e_{jt} is a mean zero random shock that is uncorrelated over time and across firms. The firm does not know the value of e_{jt} at the time it makes its decisions for period t.

OP, LP, and ACF assume that productivity follows an exogenous first-order Markov process with transition probabilities $P(\omega_{jt}|\omega_{jt-1})$. Since 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. The Bellman equation for the firm's dynamic programming problem is

$$V(k_{jt}, \omega_{jt}) = \max_{i_{jt}, r_{jt}} \pi(k_{jt}, \omega_{jt}) - c_i(i_{jt}) - c_r(r_{jt}) + \frac{1}{1+\rho} E\left[V(k_{jt+1}, \omega_{jt+1}) | k_{jt}, \omega_{jt}, i_{jt}, r_{jt}\right],$$

where $\pi(\cdot)$ denotes per-period profits and ρ is the discount rate. In the simplest case the cost functions $c_i(\cdot)$ and $c_r(\cdot)$ just transform logs into levels, but their exact forms are irrelevant for our purposes. The dynamic programming problem gives rise to two policy functions, $i(k_{jt}, \omega_{jt})$ and $r(k_{jt}, \omega_{jt})$ 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

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

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 R&D expenditures, ξ_{jt} does not: by construction ξ_{jt} is mean independent (although not necessarily fully independent) of r_{jt-1}. This *productivity innovation* represents the uncertainties that are naturally linked to productivity plus the uncertainties inherent in the R&D process (e.g., chance in discovery, degree of applicability, 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. Of course, the conditional expectation function $g(\cdot)$ is unobserved from the point of view of the econometrician (but known to the firm) and must be estimated nonparametrically.

3 Empirical strategy

Our model relaxes the assumption of an exogenous Markov process for productivity in OP and the subsequent literature. As emphasized by Ackerberg et al. (2007), endogenizing this process is problematic for the standard estimation procedures. First, it tends to invalidate the usual instrumental variables approaches. Given an exogenous Markov process, input prices are natural instruments for input quantities. This is, however, no longer the case if the transitions from current to future productivity are affected by the choice of an additional unobserved "input" such as R&D because all quantities depend on all prices. Second, in the absence of data on R&D, a critical determinant of the probability distribution of ω_{jt} given ω_{jt-1} is missing. Because both R&D and productivity are unobservable, the scalar unobservable assumption in OP is violated. This means that recovering ω_{jt} from k_{jt}, i_{jt}, and their lags, the key step in OP, may be difficult.

Buettner (2005) extends the OP approach by studying a model similar to ours while assuming transition probabilities for unobserved productivity of the form $P(\omega_{jt}|\psi_t)$, where $\psi_t = \psi(\omega_{jt-1}, r_{jt-1})$ 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. Under certain assumptions it ensures that the policy function for investment in physical capital is still invertible and that unobserved productivity can hence still be written as an unknown function of the capital stock and the investment as $\omega_{jt} = h(k_{jt}, i_{jt})$. Buettner (2005) notes, however, that there are problems with identification even when data on R&D is available.

Our estimator builds on the insight of LP that the demand for static inputs such as labor and materials can be used to recover unobserved productivity. These static inputs are chosen with current productivity known, and therefore contain information about it. Importantly, the demand for static inputs is 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. This has the advantage that our approach does not rely on an index and frees up the relationship between current productivity, R&D expenditures, and future productivity.

Our estimator differs from LP by recognizing that, given a parametric specification of the production function, the functional form of the inverse input demand functions is known. In particular, given the Cobb-Douglas production function in equation (1), the assumption that labor is a static input implies that the demand for labor is

$$l_{jt} = \frac{1}{1-\beta_l}\left(\beta_0 + \ln\beta_l + \mu + \beta_k k_{jt} + \omega_{jt} - (w_{jt} - p_{jt})\right), \tag{2}$$

where $\mu = \ln E\left[\exp(e_{jt})\right]$ and $(w_{jt} - p_{jt})$ is the real wage. Solving for ω_{jt} we obtain the inverse labor demand function

$$h(l_{jt,}k_{jt,}w_{jt} - p_{jt}) = \lambda_0 + (1 - \beta_l)l_{jt} - \beta_k k_{jt} + (w_{jt} - p_{jt}),$$

where $\lambda_0 = -\beta_0 - \ln \beta_l - \mu$ combines the constant terms. From hereon we call $h(\cdot)$ the inverse labor demand function and use h_{jt} as shorthand for its value $h(l_{jt,}k_{jt,}w_{jt} - p_{jt})$.

Substituting the inverse labor demand function $h(\cdot)$ for ω_{jt} in the production function leaves us with the marginal productivity condition for profit maximization $\ln \beta_l + (y_{jt} - l_{jt}) = w_{jt} - p_{jt} + e_{jt}$ and cancels out parameters of interest. Observing this, ACF submit that the "parametric approach does not work" (p. 16). However, using $h(\cdot)$ to substitute for ω_{jt-1} in the controlled Markov process, we have

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + g(h(l_{jt-1}, k_{jt-1}, w_{jt-1} - p_{jt-1}), r_{jt-1}) + \xi_{jt} + e_{jt}. \tag{3}$$

Both k_{jt}, whose value is determined in period $t-1$ by i_{t-1}, and r_{jt-1} are uncorrelated with ξ_{jt} by virtue of our timing assumptions. Only l_{jt} is correlated with ξ_{jt} (since ξ_{jt} is part of ω_{jt} and l_{jt} is a function of ω_{jt}). Nonlinear functions of the other variables can be used as instruments for l_{jt}, as can be lagged values of l_{jt} and the other variables. If firms can be assumed to be perfectly competitive, then the current real wage is exogenous and constitutes the most adequate instrument (since demand for labor depends directly on it).[3]

Below we discuss how imperfect competition can be taken into account and the likelihood of sample selection. Then we turn to identification and estimation. Finally, we discuss the advantages and disadvantages of our parametric inversion.

Imperfect competition. Following OP, LP, and ACF we have so far assumed a perfectly competitive environment. But when firms have some market power, say because products are differentiated, then output demand enters the specification of the inverse input demand functions (see, e.g., Jaumandreu & Mairesse 2005). Consider firms facing a downward sloping demand function that depends on the price of the output P_{jt} and the demand shifters Z_{jt}. Profit maximization requires that firms set the price that equates marginal cost to marginal revenue $P_{jt}\left(1 - \frac{1}{\eta(p_{jt}, z_{jt})}\right)$, where $\eta(\cdot)$ is the absolute value of the elasticity of demand. With firms minimizing costs, marginal cost and conditional labor demand can be determined from the cost function and combined with marginal revenue to give the

[3] Our model nests the dynamic panel model proposed by Blundell & Bond (2000). Restrict the productivity process to be exogenous and autoregressive so that we have $g(\omega_{jt-1}) = \rho\omega_{jt-1}$. Using the marginal productivity condition for profit maximization to substitute ρy_{jt-1} for $\rho(-\ln \beta_l + (w_{jt-1} - p_{jt-1}) + l_{jt-1})$, we are in the Blundell & Bond (2000) specification. Hence, the differences between their and our approach lie in the generality of the productivity process and the strategy of estimation. In the tradition of OP, LP, and ACF we replace unobserved productivity in terms of observed variables and an unpredictable component, whereas they model the same term through the use of lags of the dependent variable.

inverse labor demand function

$$h^{IC}(l_{jt}, k_{jt}, w_{jt} - p_{jt}, p_{jt}, z_{jt}) = \lambda_0 + (1-\beta_l)l_{jt} - \beta_k k_{jt} + (w_{jt} - p_{jt}) - \ln\left(1 - \frac{1}{\eta(p_{jt}, z_{jt})}\right).$$

$h^{IC}(\cdot)$ replaces $h(\cdot)$ in the estimation equation (3). As both p_{jt} and z_{jt} enter the equations lagged they are expected to be uncorrelated with the productivity innovation ξ_{jt}.[4] Of course, the current real wage can no longer be considered exogenous in an imperfectly competitive setting.

Sample selection. A potential problem in the estimation of production functions is sample selection. A firm's optimal exit decision compares the sell-off value of the firm to its expected profitability in the future. If a sufficiently adverse shock to productivity is followed immediately by exit, then there is a negative correlation between the shocks and the capital stocks of the firms that remain in the industry, and biased estimates result.

Accounting for R&D expenditures in the Markov process tends to alleviate the selection problem. Innovative activities often imply large sunk cost that make a firm more reluctant to exit the industry or at least to exit it immediately. Moreover, the firm now has an instrument to try to rectify an adverse shock. Given that the institutional setting in Spain renders it further unlikely that a firm is able to exit the industry immediately after receiving an adverse shock to productivity (see, e.g., Djankov, La Porta, Lopez-de Silanes & Shleifer 2002), we follow LP and ACF and do not correct for sample selection, although this could be done by modeling exit decisions along the lines of OP. At this stage we simply explore whether there is a link between exit decision and estimated productivity (see Section 5.1 for details).

3.1 Identification

We show that our estimation equation (3) identifies the parameters of the production function and the Markov process that governs the evolution of productivity. We first study a setting with an exogenous Markov process as familiar from OP, LP, and ACF. We then turn to the controlled Markov process that accounts for the impact of R&D on productivity.

Exogenous Markov process. With an exogenous Markov process the estimation equation is similar to equation (3) except that the conditional expectation function $g(\cdot)$ does not depend on r_{jt-1}. This equation is a semiparametric, so-called partially-linear, model with the additional restriction that the inverse labor demand function $h(\cdot)$ is of known form.

To see how the known form of $h(\cdot)$ aids identification, suppose to the contrary that $h(\cdot)$ is of unknown form. In this case, the composition of $h(\cdot)$ and $g(\cdot)$ is another function of unknown form. The fundamental condition for identification is that the variables in the

[4]Note that this setting yields an estimate of the average elasticity of demand. This is possible for the same reason that correcting the Solow residual for imperfect competition allows for estimating margins and elasticities (see, e.g., Hall 1990).

parametric part of the model are not perfectly predictable (in the least squares sense) by the variables in the nonparametric part (Robinson 1988). In other words, there cannot be a functional relationship between the variables in the parametric and nonparametric parts (see Newey, Powell & Vella 1999).

The fundamental identification condition is violated if $h(\cdot)$ is of unknown form because k_{jt} is perfectly predictable from the variables in the nonparametric part. Hence, there is a collinearity problem similar to the one that ACF ascertain for the estimator in LP. To see this, recall that $K_{jt} = (1-\delta)K_{jt-1} + \exp(i(k_{jt-1}, \omega_{jt-1}))$ by the law of motion and the policy function for investment in physical capital. But k_{jt-1} is one of the arguments of $h(\cdot)$ and ω_{jt-1} is by construction a function of all arguments of $h(\cdot)$; thus, k_{jt} can be inferred from k_{jt-1} and $i(k_{jt-1}, \omega_{jt-1})$.

Our approach differs from LP in that it exploits the known form of the inverse labor demand function $h(\cdot)$. In this case, the central question becomes whether k_{jt} is perfectly predictable from h_{jt-1}, the value of $h(\cdot)$, as opposed to the arguments of $h(\cdot)$. This is not the case: While h_{jt-1} is identical to ω_{jt-1}, we cannot infer k_{jt-1} from h_{jt-1}. Since there is no functional relationship between k_{jt} in the parametric part and h_{jt-1} in the nonparametric part, the model is identified. Note that firm-level wages and prices do not contribute to identification.

In sum, by using a parametric rather than nonparametric inversion we avoid a collinearity problem similar to the one in LP. If the productivity process can indeed be taken as exogenous, then our estimator offers an alternative to the estimator proposed by ACF. We further discuss the advantages and disadvantages of our estimator in Section 3.3.

Controlled Markov process. With a controlled Markov process the conditional expectation function $g(\cdot)$ in our estimation equation (3) depends on r_{jt-1} in addition to ω_{jt-1}. Because $h(\cdot)$ is of known form, we have to ask if k_{jt} in the parametric part can be inferred from h_{jt-1} and r_{jt-1} in the nonparametric part. This may indeed be possible. Recall that $r_{jt-1} = r(k_{jt-1}, \omega_{jt-1})$ by the policy function for investment in knowledge. Hence, if its R&D expenditures happen to be increasing in the capital stock of the firm, then $r(\cdot)$ can be inverted to back out k_{jt-1}. In this case, k_{jt} can be inferred from k_{jt-1} and $i(k_{jt-1}, \omega_{jt-1})$, and the model is not identified.

Fortunately, there is little reason to believe that this is the case. In fact, even under the fairly stringent assumptions in Buettner (2005), it is not clear that $r(\cdot)$ is invertible. Moreover, there is empirical evidence that invertibility may fail even for investment in physical capital (Greenstreet 2005) and it seems clear that R&D expenditures are even more fickle.

Even if $r(\cdot)$ happens to be an invertible function of k_{jt-1}, anything that shifts the costs of the investments in physical capital and knowledge over time guarantees identification. For example, investment opportunities and the price of equipment goods are likely to vary and the marginal cost of investment in knowledge depends greatly on the nature of the undertaken project. Using x_{jt} to denote these shifters, the policy functions become $i(k_{jt}, \omega_{jt}, x_{jt})$ and $r(k_{jt}, \omega_{jt}, x_{jt})$. Obviously, x_{jt} cannot be perfectly predicted from h_{jt-1} and r_{jt-1}. This breaks the functional relationship between $K_{jt} = (1-\delta)K_{jt-1} + \exp(i(k_{jt-1}, \omega_{jt-1}, x_{jt}))$ and h_{jt-1} and r_{jt-1}.[5] Note that the model is identified even if x_{jt} is unobserved.

3.2 Estimation

The estimation problem can be cast in the nonlinear GMM framework

$$E\left[z'_{jt}(\xi_{jt} + e_{jt})\right] = E\left[z'_{jt} v_{jt}(\theta)\right] = 0,$$

where z_{jt} is a vector of instruments and we write the error term $v_{jt}(\cdot)$ as a function of the parameters θ to be estimated. The objective function is

$$\min_{\theta} \left[\frac{1}{N} \sum_j z'_j v_j(\theta)\right]' A_N \left[\frac{1}{N} \sum_j z'_j v_j(\theta)\right],$$

where z'_j and $v_j(\cdot)$ are $L \times T_j$ and $T_j \times 1$ vectors, respectively, with L being the number of instruments, T_j being the number of observations of firm j, and N the number of firms. We first use the weighting matrix $A_N = \left(\frac{1}{N}\sum_j z'_j z_j\right)^{-1}$ to obtain a consistent estimator of θ and then we compute the optimal estimator with weighting matrix $A_N = \left(\frac{1}{N}\sum_j z'_j v_j(\hat{\theta}) v_j(\hat{\theta})' z_j\right)^{-1}$.

Production function and Markov process. While we have so far assumed a value-added production function in order to simplify the exposition, from hereon we assume a gross-output production function. We further assume imperfect competition. Our preliminary estimates indicate that in some industries it is useful to add a time trend or dummies to the production function. In case of a time trend our goal is to estimate

$$y_{jt} = \beta_0 + \beta_t t + \beta_l l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + g(h^{IC}_{jt-1} - \lambda_0, r_{jt-1}) + \xi_{jt} + e_{jt},$$

where

$$h^{IC}_{jt} - \lambda_0 = -\beta_t t + (1-\beta_l-\beta_m) l_{jt} - \beta_k k_{jt} + (1-\beta_m)(w_{jt}-p_{jt}) + \beta_m(p_{Mjt}-p_{jt}) - \ln\left(1 - \frac{1}{\eta(p_{jt}, z_{jt})}\right)$$

and $(p_{Mjt} - p_{jt})$ is the real price of materials. The case of dummies is analogous. Note that the constant of the unknown function $g(\cdot)$ subsumes any constant that its arguments may

[5] Identification is also restored if an error term is added to the law of motion for physical capital so that k_{jt} can no longer be written as a function of h_{jt-1} and r_{jt-1}. Accounting for uncertainty in the impact of investment may be appropriate depending on the construction of the capital stock in the data.

have. 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 $g(h^{IC}_{jt-1}-\lambda_0, r_{jt-1})$ as

$$1(R_{jt-1}=0)\left(g_{00}+g_{01}(h^{IC}_{jt-1}-\lambda_0)\right)+1(R_{jt-1}>0)\left(g_{10}+g_{11}(h^{IC}_{jt-1}-\lambda_0, r_{jt-1})\right).$$

Since the constants g_{00}, g_{10}, and β_0 cannot be separately estimated, we jointly estimate β_0+g_{00} and include a dummy for performers to measure $\beta_0+g_{10}-(\beta_0+g_{00})=g_{10}-g_{00}$.

Series estimator. We model an unknown function $q(\cdot)$ of one variable v by a univariate polynomial of degree Q. We model an unknown function $q(\cdot)$ of two variables v and u by a complete set of polynomials of degree Q (see Judd 1998). In the remainder of this paper we set $Q=3$. We specify the absolute value of the elasticity of demand as $1+\exp(q_0+q(p_{jt-1}, z_{jt-1}))$ in order to impose the theoretical restriction $\eta(p_{jt-1}, z_{jt-1})>1$.

Instrumental variables. 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 $\eta(\cdot)$.

As discussed before, k_{jt} is always a valid instrument because it is not correlated with ξ_{jt} as the latter is unpredictable when i_{t-1} is chosen.[6] Labor and materials, however, are contemporaneously correlated with the innovation to productivity. While the lags of both these inputs are valid instruments, l_{jt-1} is already appearing in h^{IC}_{jt-1}. We can still use m_{jt-1}. Constant and time trend are valid instruments. We therefore have four instruments to estimate the constant and the coefficients for the time trend, capital, labor, and materials. This leaves us short of at least one more instrument.

Following Wooldridge (2004) we use as additional instruments the complete set of polynomials of degree three in the variables l_{jt-1}, k_{jt-1}, $w_{jt-1}-p_{jt-1}$, and $p_{Mjt-1}-p_{jt-1}$ (34 instruments) as well as the complete set of polynomials in the variables p_{jt-1} and z_{jt-1} (9 instruments). We further use the powers up to degree three in the variable r_{jt-1} (3 instruments) and the interactions up to degree three of it with the above two complete sets of polynomials ($12+6=18$ instruments). We finally use a dummy for the firms that perform R&D. This gives a total of 69 instruments.

When we have enough degrees of freedom we interact the complete set of polynomials in the variables l_{jt-1}, k_{jt-1}, $w_{jt-1}-p_{jt-1}$, and $p_{Mjt-1}-p_{jt-1}$ as well as the complete set of polynomials in the variables p_{jt-1} and z_{jt-1} with dummies for nonperformers and performers for a total of $69+34+9=112$ instruments.

Given these instruments, our estimator is the GMM version of Ai & Chen's (2003)

[6] 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.

sieve minimum distance estimator, a nonparametric least squares technique (see Newey & Powell 2003). Hence, even if the variables in the conditional expectation function $g(\cdot)$ are correlated with the error term of the estimation equation, we can still obtain consistent estimates by restricting the instruments to the exogenous conditioning variables.

Productivity estimates. Once the model is estimated we can recover actual productivity $\omega_{jt} = h^{IC}_{jt}$ and expected productivity $g(\cdot)$ up to a constant. We estimate the productivity innovation ξ_{jt} up to a constant as the difference between ω_{jt} and $g(\cdot)$. We can also estimate the random shocks e_{jt}. In the remainder of this paper we let $\widehat{\omega}_{jt} = \widehat{h}^{IC}_{jt}$, $\widehat{g}(\cdot)$, $\widehat{\xi}_{jt}$, and $\widehat{e}_{jt}$ denote these estimates.

3.3 Parametric vs. nonparametric inversion

OP, LP, and ACF use nonparametric methods to recover unobserved productivity. This forces them either to rely on a two-stage procedure or to jointly estimate a system of equations as suggested by Wooldridge (2004). The drawback of the two-stage approach is a loss of efficiency and the additional effort required to compute standard errors. In contrast, the joint estimation of a system of equations is numerically more demanding (as it requires a nonlinear optimization over a larger set of parameters).[7]

We differ from the previous literature in that we recognize that the parametric specification of the production function in combination with the assumption that labor is a static input implies a known form for the inverse labor demand function $h(\cdot)$ that can be used to control for unobserved productivity. Making full use of the structural assumptions aids identification (Section 3.1). Moreover, we have but a single equation to estimate, and only the conditional expectation function $g(\cdot)$ is unknown and must be estimated nonparametrically (Section 3.2). The former eases the computational burden and the latter yields efficiency gains. These efficiency gains may be especially important in cases where one has a limited amount of data or the inverse labor demand function $h(\cdot)$ has a large number of arguments.

A seeming drawback of our parametric approach is that it requires firm-level wage and price data.[8] However, the same is true for a nonparametric approach: The demand for labor is a function of the real wage whether one inverts it parametrically or nonparametrically; by spelling out the demand for labor in equation (2) our parametric approach just makes the role of the real wage explicit.

In the absence of firm-level wage and price data, one may be able to assume that the industry is perfectly competitive and replace the real wage by a set of dummies (as in LP

[7] See Ackerberg et al. (2007) for a discussion of the relative merits of the two approaches.

[8] In our data wages and prices vary both across firms and across periods. This variation is partly due to geographic and temporal differences in the supply of labor and the fact that firms operate in different product submarkets. But even absent exogenous variation our model of imperfect competition implies that the real wage varies with a firm's position in the product market.

and ACF) or aggregate wage and price indices. But in this case one may have to confront an issue raised by Bond & Söderbom (2005). They argue that, absent any variation in prices, it may be hard to estimate the coefficients on static inputs in a Cobb-Douglas production function.

At a very general level there is a tradeoff between the additional efficiency of a parametric approach and the additional robustness of a nonparametric approach along the lines of OP, LP, and ACF. Our parametric approach rests on the assumption that the demand for labor is the solution to the firm's short-run profit maximization problem. By fully exploiting this assumption we are able to devise a particularly simple estimator for production functions. However, this assumption may or may not be valid in a given application, and in Section 5.1 we provide a formal statistical test for it. The advantage of a nonparametric approach is that it can accommodate situations where this assumption is violated and the labor decision has dynamic consequences. The disadvantage is that it also relies on specific assumptions that may or may not be valid in a given application. The estimator proposed by ACF, in particular, assumes that labor is chosen after capital but before materials are chosen, that productivity evolves in the interim, and that none of the inputs that appear in the production function are chosen in accordance with the firm's short-run profit maximization problem. The parametric and nonparametric "fixes" to the collinearity problem in LP are therefore very much complementary.

4 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 all firms of this size chose to respond. Some firms vanish from the sample, due to both exit (shutdown by death or abandonment of activity) and 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.

Given that our estimation procedure requires a lag of one year, we restrict the sample to firms with at least two years of data. The resulting sample covers a total of 1879 firms. Columns (1) and (2) of Table 1 show the number of observations and firms by industry. The samples are of moderate size. 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. The growth of prices, averaged from the growth of prices as reported individually by each firm, is moderate (column (9)).

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).[9] The manufacturing sector consists partly of transnational companies with production facilities in Spain and huge 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.[10,11]

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 1.1% 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%.

5 Estimation results

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

[9] 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).

[10] 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, Jaumandreu & Pazo 2005). This suggests that R&D expenditures irrespective of their origin are the relevant variable for explaining productivity.

[11] 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.

productivity is subject to a high degree of nonlinearity and uncertainty. Then we provide a more detailed comparison between our model of endogenous productivity change and the knowledge capital model and show that the data refute the assumptions at the heart of the knowledge capital model.

Our model is richer than the knowledge capital model, in particular with regard to the treatment of heterogeneity. In order 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.

5.1 Production function and Markov process

Table 2 summarizes different production function estimates.[12] Columns (1)–(3) report the coefficients estimated from OLS regressions of the log of output on the logs of inputs. The coefficients are reasonable and returns to scale, as given by $\beta_l + \beta_k + \beta_m$, are close to constant. The share of capital in value added, as given by $\frac{\beta_l}{\beta_l+\beta_k}$, is between 0.15 and 0.35 as expected.

Columns (4)–(9) of Table 2 report the coefficients estimated when we use the demand for labor to back out unobserved productivity. Treating labor as a static input is appropriate because Spain greatly enhanced the possibilities for hiring and firing temporary workers during the 1980s. By the beginning of the 1990s the share of temporary workers in the manufacturing sector had stabilized in excess of a quarter, one of the highest shares in Europe. Rapid expansion and contraction of the number of temporary workers became common (Dolado, Garcia-Serrano & Jimeno 2002). In addition, we measure labor 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 labor.[13]

Columns (4)–(6) pertain to an exogenous Markov process. Compared to the OLS regressions, the changes go in the direction expected from theory and match the results in OP and LP. The labor coefficients decrease considerably in 8 industries while the capital coefficients increase somewhat in 6 industries. The materials coefficients show no particular pattern.

[12] The specification of the production function for industries 2, 3, 6, and 10 includes a time trend and dummies for industries 1, 7, and 8. We use the larger set of instruments for industries 1, 3, 6, 7, and 8. The demand shifter z_{jt} is an index of market dynamism (see Appendix A for details).

[13] The results broadly agree when we use the demand for materials to back out unobserved productivity. One notable difference is that the materials (labor) coefficient in the production function tends to be lower when we use the demand for materials (labor). The most likely explanation is that the term in the used input that appears directly in the production function and the terms that appear inside the conditional expectation function are slightly collinear. This may make it harder to precisely estimate the coefficient of the used input and suggests combining the inverse input demand functions to estimate the production function parameters. How to do this in a manner that is consistent with the modeling framework is a topic for future research.

Columns (7)–(9) pertain a controlled Markov process. Again, compared to the OLS regressions, the changes go in the expected direction. The labor coefficients decrease, the capital coefficients increase in 7 cases and are virtually unchanged in 2 more cases. Specifying a controlled instead of an exogenous Markov process tends to further decrease the labor coefficient and increase the capital coefficient. This leaves open the question whether it is possible to obtain consistent estimates of the parameters of the production function in the absence of data on R&D, although it is clear that omitting R&D expenditures from the Markov process substantially distorts the retrieved productivities (see Section 5.4 for details).

Imperfect competition. We test for perfect competition by removing the function in the equilibrium price p_{jt} and the demand shifter z_{jt} from $h^{IC}(\cdot)$ in order to obtain $h(\cdot)$.[14] The data very clearly reject perfect competition, see columns (1) and (2) of Table 3. Our estimates of the average elasticity of demand are around 2 (column (3)).

Specification tests. To assess the validity of our estimates we have conducted a series of tests as reported in columns (4)–(11) of Table 3. We first test for overidentifying restrictions or validity of the moment conditions.[15] With the exception of industry 1, the validity of the moment conditions cannot be rejected by a wide margin, see columns (4) and (5).

Lagged labor and lagged 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 when all moments are included to its value when the moments involving either lagged labor or lagged materials are excluded. As columns (6)–(9) of Table 3 show, the validity of lagged labor and lagged materials as instruments cannot be rejected. We also test the subset of moments involving capital and lagged capital. As columns (10) and (11) of Table 3 show, the exogeneity assumption on capital and lagged capital is rejected at the usual significance levels for industries 1, 2, and 3. When viewed in conjunction with the other tests, however, we feel that there is little ground for concern regarding the validity of the moment conditions in these industries.

Taken together, the above tests also support our choice of the functional form for the production function: Had the assumed linearity in the log of inputs been violated, then at least part of the nonlinearity would have been pushed into the productivity innovation, thereby increasing the test statistics.

[14]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.

[15]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.

Parametric inversion. Our final test validates directly our parametric inversion. Recall that the production function parameters appear both in the production function and in the inverse labor demand function. If the inverse labor demand function were misspecified, say because the labor decision has dynamic consequences, then this would cause β_l and β_k in the inverse labor demand function to diverge from their counterparts in the production function. We therefore test whether the structural parameters in the two parts of the model are equal. As columns (12) and (13) of Table 3 show there is no evidence that our parametric inversion is inappropriate, although this is somewhat less clearcut in industries 6 and 7.

Sample selection. To assess the potential for sample selection we explore whether there is a link between exit decisions and estimated productivity. After controlling for year and industry effects it appears that the productivity of exitors is lower on average than that of incumbents one year before exiting in case of small firms and two years before exiting in case of large firms. Moreover, this drop in productivity is more pronounced for firms that engage in R&D than for those that do not. We cannot reject the null hypothesis that exitors are systematically less productive than incumbents in all cases but one, although this may be due to the small number of exitors in our sample. Taken together, this suggests that firms tend to exit the industry sometime after receiving an adverse shock to productivity, thereby decreasing the scope for sample selection.

The productivity of entrants appears to be lower on average than that of incumbents, similar to earlier results in Baily, Hulten & Campbell (1992) and Huergo & Jaumandreu (2004). We reject the null hypothesis that entrants are systematically more productive than incumbents in all cases but one. In addition, there is some evidence that entrants that engage in R&D are quicker in catching up with incumbents than those that do not.

5.2 Nonlinearity and uncertainty

A key contribution of this paper is to endogenize the productivity process. We assess the role of R&D by comparing the controlled with the exogenous Markov process. To this end, we test whether R&D can be excluded from the conditional expectation function for performers and whether the common part of the conditional expectation functions for performers and nonperformers are equal, i.e., whether $g_{11}(h^{IC}_{jt-1} - \lambda_0, r_{jt-1}) = g_{01}(h^{IC}_{jt-1} - \lambda_0)$ for all r_{jt-1}. As columns (1) and (2) of Table 4 shows, the result is overwhelming: In all cases the constraints imposed by the model with the exogenous Markov process are clearly rejected.

To get a better sense of the importance of R&D we use a standard growth decomposition. Roughly two thirds of the growth in output is explained by the growth in inputs, with the glaring exception of industry 8 where output is growing while inputs are shrinking. While there are considerable differences across industries, about one half of the year-to-year variation in expected productivity is due the variation in R&D expenditures. While these numbers already hint at the major role played by R&D, they have to be interpreted as lower bounds because a part of the impact of current R&D expenditures persists and is carried forward into future productivity. We will come back to the persistence in productivity in Section 5.6.

Nonlinearity. As a first step in exploring the link between R&D and productivity, we test whether the conditional expectation function $g(\cdot)$ is separable in current productivity and R&D expenditures, i.e., whether $g_{11}(h_{jt-1}^{IC} - \lambda_0, r_{jt-1})$ for firms that perform R&D can be broken up into two additively separable functions $g_{11}(h_{jt-1}^{IC} - \lambda_0)$ and $g_{12}(r_{jt-1})$. Columns (3) and (4) of Table 4 indicate that 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.

We further illustrate the economic significance of these interactions in columns (5)–(8) of Table 4. We list 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) so that current productivity and (the level of) R&D expenditures are, at least locally, complements (substitutes) in the accumulation of productivity. There is evidence of complementarities in industries 1, 3, 9, and 10 and to some extent also in industries 4 and 6. We also list the percentage of observations where $\frac{\partial^2 g(\omega_{jt-1}, r_{jt-1})}{\partial R_{jt-1}^2} = \frac{1}{R_{jt-1}^2} \left(\frac{\partial^2 g(\omega_{jt-1}, r_{jt-1})}{\partial r_{jt-1}^2} - \frac{\partial g(\omega_{jt-1}, r_{jt-1})}{\partial r_{jt-1}} \right)$ is significantly positive (negative) so that there are locally increasing (decreasing) returns to R&D. There is evidence of increasing returns to R&D in industries 3, 6, 7, 8, 9 and 10 (see also Klette 1996). At the same time, however, a substantial fraction of firms seems to operate under decreasing returns to R&D in industries 3, 4, 6, 7, and 9.

Uncertainty. Turning from nonlinearity to uncertainty, column (9) of Table 4 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 (10) of Table 4 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 attained productivity, between 20% and 60%. 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 the least R&D intensive industries. We will come back to the economic significance of the uncertainties inherent in the R&D process in Section 5.7.

5.3 Knowledge capital model

The knowledge capital model of Griliches (1979) has remained a cornerstone of the productivity literature. While it has been used in hundreds of studies on firm-level productivity, the underlying empirical strategy has changed little over the years (see the surveys in Mairesse & Sassenou (1991), Griliches (1995), and Griliches (2000)). The most widely used form of the knowledge capital model adds the log of knowledge capital c_{jt} as an extra input to a Cobb-Douglas production function yielding

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + \varepsilon c_{jt} + e_{jt}, \tag{4}$$

where ε is the elasticity of output with respect to knowledge capital. Knowledge capital is assumed to accumulate with R&D expenditures and to depreciate from period to period at a rate δ. Hence, its law of motion can be written as $C_{jt} = (1-\delta)C_{jt-1} + R_{jt-1}$.

Estimating equation (4) requires constructing C_{jt}, the stock of knowledge capital of firm j in period t, from observed R&D expenditures. There are two obvious problems: First, we have little idea what the appropriate rate of depreciation δ is.[16] Second, our history of R&D expenditures is often not very long. Following Hall & Mairesse (1995) we therefore assume that the rate of depreciation is 0.15 per period and estimate the initial capital from the date of birth of the firm by extrapolating its average R&D expenditures during the time that it is observed.

Replacing εc_{jt} by ω_{jt} in equation (4) shows that the basic form of the knowledge capital model can be understood as a special case of our dynamic investment model in which the stochastic process that governs productivity has degenerated to a deterministic process. We test this model against our model using a nonnested test.[17] Given that we have already shown that nonlinearity and uncertainty play a large role in the link between R&D and productivity, it does not come as a surprise that the data very clearly reject the knowledge capital model in its basic form, see columns (1) and (2) of Table 5.

We next develop and evaluate two generalizations of the knowledge capital model. Our

[16] Estimating the rate of depreciation δ using distributed lag models is notoriously difficult, even in case of physical capital (Pakes & Schankerman 1984*b*, Pakes & Griliches 1984, Nadiri & Prucha 1996). Our own attempts at estimating δ together with the parameters in equation (4) have largely been unsuccessful.

[17] We apply the Rivers & Vuong (2002) test for model selection among nonnested models. The difference between the GMM objective functions, scaled by $\sqrt{N}$, has an asymptotic normal distribution with variance

$$\begin{aligned}\sigma^2 = \; & 4\Big[(\sum_j z_j' v_j(\hat\theta))' A_N (\sum_j z_j' v_j(\hat\theta) v_j(\hat\theta)' z_j) A_N (\sum_j z_j' v_j(\hat\theta)) \\ & + (\sum_j z_j' v_j(\hat\theta^{KCM}))' A_N (\sum_j z_j' v_j(\hat\theta^{KCM}) v_j(\hat\theta^{KCM})' z_j) A_N (\sum_j z_j' v_j(\hat\theta^{KCM})) \\ & - 2(\sum_j z_j' v_j(\hat\theta))' A_N (\sum_j z_j' v_j(\hat\theta) v_j(\hat\theta^{KCM})' z_j) A_N (\sum_j z_j' v_j(\hat\theta^{KCM}))\Big],\end{aligned}$$

where $\hat\theta$ and $\hat\theta^{KCM}$ are the unrestricted and restricted parameter estimates, respectively, the instruments in z_j are kept the same, and A_N is a common first-step weighting matrix.

first generalization is to add unobserved productivity ω_{jt} back into equation (4) yielding

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

We follow OP, LP, and ACF and assume that productivity follows an exogenous Markov process. Note that this model has two Markov processes, a deterministic one for the stock of knowledge capital and a stochastic one for productivity. Because the sum of two Markov processes is not necessarily a Markov process, this model is not nested in our model of endogenous productivity change with just one Markov process. Again we use a nonnested test. The data strongly favor our model, see columns (3) and (4) of Table 5.

The basic form of the knowledge capital model abstracts from uncertainty in the R&D process.[18] Our second generalization aims to address this issue. To capture that the accumulation of improvements to productivity is likely to be subjected to shocks, we borrow from the dynamic investment model of Hall & Hayashi (1989) and assume that the law of motion is $C_{jt} = (1-\delta)C_{jt-1} + R_{jt-1} + \frac{1}{\varepsilon}C_{jt-1}\xi_{jt} = C_{jt-1}\left(1 - \delta + \frac{R_{jt-1}}{C_{jt-1}} + \frac{1}{\varepsilon}\xi_{jt}\right)$ so that the effect of the rate of investment in knowledge $\frac{R_{jt-1}}{C_{jt-1}}$ has an unpredictable component ξ_{jt}. Note that it is no longer possible to construct the stock of knowledge capital from observed R&D expenditures and, with it, control for the impact of R&D on productivity.

To make this model estimable we recast it within our setting. Taking logs and letting $\omega_{jt} = \varepsilon c_{jt}$ the law of motion can be written as

$$\omega_{jt} \simeq \omega_{jt-1} + \varepsilon\left(\frac{\exp(r_{jt-1})}{\exp(\omega_{jt-1}/\varepsilon)} - \delta\right) + \xi_{jt} \tag{5}$$

and hence $\omega_{jt} = g(\omega_{jt-1}, r_{jt-1}) + \xi_{jt}$. This turns out to be our controlled first-order Markov process but with a particular functional form for the conditional expectation function $g(\cdot)$. We use once again a nonnested test. The data very clearly reject the functional form restrictions implied by the knowledge capital model, see columns (5) and (6) of Table 5.[19]

In sum, the linearity and certainty assumptions in the accumulation and depreciation of knowledge that underpin the knowledge capital model may have to be relaxed in order to properly assess the impact of the investment in knowledge on the productivity of firms. The unspecified form of the law of motion and the random nature of accumulation in our setting are important advantages over the knowledge capital model. Moreover, by treating productivity as unobserved, we circumvent the initial conditions problem in the knowledge capital model. Finally, our empirical strategy takes into account that the endogeneity problem in production function estimation may not be completely resolved by adding the stock of knowledge capital to the conventional factors of production.

[18] Griliches (1979) allows for a random shock to the constructed stock of knowledge capital (see also Griliches 1998). Being entirely transitory, however, this shock is absorbed by the error term of the production function and hence does not capture uncertainty in the R&D process.

[19] We continue to reject when we base the test on the exact form for the law of motion implied by the knowledge capital model rather than the approximate form in equation (5).

5.4 Productivity levels

To describe differences in expected productivity between firms that perform R&D and firms that do not perform R&D, we employ kernels to estimate the density and the distribution functions associated with the subsamples of observations with R&D and without R&D. To be able to interpret these and other descriptive measures in the remainder of the paper as representative aggregates, we proceed as described in Section 4. Figure 1 shows the density and distribution functions for performers (solid line) and nonperformers (dashed line) for each industry. In most industries the distribution for performers is to the right of the distribution for nonperformers. This strongly suggests stochastic dominance. In contrast, the distribution functions cross in industries 2 and 4 as well as in industries 9 and 10 that exhibit low innovative activity. This violation of stochastic dominance is mild in case of industry 2 whereas in industry 4 attaining the highest productivity levels is clearly more likely for the nonperformers than for the performers.

Before formally comparing the means and variances of the distributions and the distributions themselves, we illustrate the impact of omitting R&D expenditures from the Markov process of unobserved productivity. We have added the so-obtained density and distribution functions to Figure 1 (dotted line). Comparing them to the density and distribution functions for a controlled Markov process reveals that the exogenous process takes a sort of average over firms with distinct innovative activities and hence blurs remarkable differences in the impact of the investment in knowledge on the productivity of firms.

Mean and variance. Turning to the moments of the distributions, the difference in means is computed as

$$\begin{aligned}\widehat{\overline{g}}_0 - \widehat{\overline{g}}_1 \quad = \quad & \frac{1}{NT_0} \sum\nolimits_j \sum\nolimits_t 1(R_{jt-1} = 0) \widehat{g}_{01}(\widehat{h}^{IC}_{jt-1}) \\ & - \frac{1}{NT_1} \sum\nolimits_j \sum\nolimits_t 1(R_{jt-1} > 0)[(\widehat{g_{10} - g_{00}}) + \widehat{g}_{11}(\widehat{h}^{IC}_{jt-1}, r_{t-1})],\end{aligned}$$

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

$$t = \frac{\widehat{\overline{g}}_0 - \widehat{\overline{g}}_1}{\sqrt{Var(g_{01})/(NT_0 - 1) + Var(g_{11})/(NT_1 - 1)}}$$

which follows a t distribution with $\min(NT_0, NT_1) - 1$ degrees of freedom and the variances using

$$F = \frac{Var(g_{01})}{Var(g_{11})}$$

which follows an F distribution with $NT_0 - 1$ and $NT_1 - 1$ degrees of freedom. To account for the survey design we conduct separate tests on the subsamples of small and large firms.

Column (4) of Table 6 reports the difference in means $\widehat{\overline{g}}_1 - \widehat{\overline{g}}_0$ (with the opposite sign of the test statistic for the sake of intuition) and columns (5)–(8) report the standard deviations and the test statistics along with their probability values separately for the subsamples of small and large firms. 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. The differences are sizable, with many values 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, at the usual significance levels, the formal statistical test rejects the hypothesis of a higher mean of expected productivity among performers than among nonperformers only in case of large firms in industry 4.

The hypothesis of greater variability for performers than for nonperformers is rejected in many cases, although there does not seem to be a recognizable pattern. As can be seen in columns (9) and (10) of Table 6, it is rejected for both size groups in industries 6 and 7, for small firms in industry 3, and for large firms in industries 1, 2, and 4.

Distribution. The above results suggest to compare the distributions themselves. We use a Kolmogorov-Smirnov test to compare the empirical distributions of two independent samples (see Barret & Donald (2003) and Delgado, Farinas & Ruano (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, where for occasional performers we average only over the years with R&D (and discard the years without R&D). This avoids dependent observations and sets the sample sizes equal to the number of nonperformers and performers, N_0 and N_1, respectively.

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

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

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 test tends to be inconclusive when the number of firms is small, we limit it to cases in which we have at least 20 performers and 20 nonperformers. The results are reported in columns (11)–(14) of Table 6. Equality of distributions is rejected in six out of ten cases. Stochastic dominance cannot be rejected anywhere.

To further illustrate the consequences of omitting R&D expenditures from the Markov process of unobserved productivity, we have redone the above tests for the case of an exogenous Markov process. The results are striking: We can no longer reject the equality of the productivity distributions of performers and nonperformers in four out of the six cases where we rejected the equality of the productivity distributions for a controlled Markov process. The upper panels of Figure 2 show at the example of industry 6 that the density and distribution functions for performers (solid line) and nonperformers (dashed line) are virtually indistinguishable if an exogenous Markov process is assumed. As can be seen in the lower panels of Figure 2, the same happens if we use an alternative estimator such as OP that assumes that productivity is exogenous. This once more makes apparent that omitting R&D expenditures substantially distorts the retrieved productivities.

In sum, comparing expected productivity across firms that perform R&D and firms that do not perform R&D we find strong evidence of stochastic dominance in most industries. It remains to be explained how the expected productivity can possibly be lower for performers than for nonperformers in some industries. One explanation is the heterogeneity across the firms within an industries that arises partly because firms engage in R&D to various degrees and partly because the level of aggregation used in defining these industries encompasses many different specific innovative activities. Stochastic dominance may hold if we were able to split these industries into more homogeneous innovative activities.

5.5 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}, we compute the expectation of productivity growth as

$$\beta_t + E(\omega_{jt} - \omega_{jt-1} | \omega_{jt-1}, r_{jt-1,}) = \beta_t + g(\omega_{jt-1}, r_{jt-1}) - \omega_{jt-1}. \tag{6}$$

Using the fact that the innovation to productivity has mean zero, i.e., $E(\xi_{jt-1}|\omega_{jt-2}, r_{jt-2}) = 0$, we estimate the average of the expectation of productivity growth as $\beta_t + \frac{1}{N}\sum_j \sum_t \frac{1}{T_j}[\widehat{g}(\widehat{h}^{IC}_{jt-1}, r_{jt-1}) - \widehat{g}(\widehat{h}^{IC}_{jt-2}, r_{jt-2})]$. To account for the survey design we replicate the subsample of small firms $\frac{70}{5} = 14$ times before merging it with the subsample of large firms. Columns (1)–(3) of Table 7 report the results for the entire sample and for the subsamples of observations with and without R&D. We also compute a weighted version to be able to interpret the expectation of productivity growth as representative for an industry as a whole. 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. Assuming that $E(\mu_{jt}\xi_{jt-1}|\omega_{jt-2}, r_{jt-2}) = 0$, we estimate the average as $\beta_t + \frac{1}{T}\sum_t \sum_j \mu_{jt}[\widehat{g}(\widehat{h}^{IC}_{jt-1}, r_{jt-1}) - \widehat{g}(\widehat{h}^{IC}_{jt-2}, r_{jt-2})]$. Columns (4)–(6) of Table 7 report the results along with a decomposition into the contributions of observations with and without R&D.

Productivity growth is higher for performers than for nonperformers in 6 industries, sometimes considerably so. Taken together these industries account for over two thirds of manufacturing output (see Appendix A for details). The industries in which the relationship is reversed are again industries 4 and 10 and now also industry 8. The standard deviations indicate that there are considerable differences in productivity growth within firms that engage in R&D as well as within those that do not. Productivity growth is more variable for performers than for nonperformers in five out of nine industries, including industries 8 and 10. This indicates that the productivity of at least some performers tends to grow much faster than the productivity of nonperformers, even though on average performers exhibit slower productivity growth than nonperformers in these industries.

A comparison of unweighted and weighted productivity growth shows that there is no definite pattern in productivity growth by size group: The productivity of small firms grows more rapidly in some industries and less in others. What is clear, however, is that productivity growth is highest in some of the industries with high innovative activity (above 2% in industry 3, above 1.5% in industry 4, and above 3% in industry 6) followed by some of the industries with intermediate innovative activity (above 1.5% in industry 1).

Columns (5) and (6) of Table 7 are particularly important. The contribution to productivity growth of firms that perform R&D is estimated to explain between 65% and 80% of productivity growth in the industries with high innovative activity and between 70% and 85% in the industries with intermediate innovative activity (with the exception of industry 8). This is all the more remarkable since in these industries between 35% and 45% and between 10% and 20% of firms engage in R&D. While these firms manufacture between 70% and 75% of output in the industries with high innovative activity and between 45% and 55% in the industries with intermediate innovative activity, their contribution to productivity growth exceeds their share of output by between 5% and 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. R&D expenditures are thus indeed a primary source of productivity growth.

5.6 Return to R&D and persistence in productivity

How hard must a firm work to maintain and advance its productivity? Recall that a change in the conditional expectation function $g(\cdot)$ can be interpreted as the expected percentage change in total factor productivity. Hence, $\frac{\partial g(\omega_{jt-1},r_{jt-1})}{\partial r_{jt-1}}$ is the elasticity of output with respect to R&D expenditures or a measure of the return to R&D. 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.

Columns (1)–(4) of Table 8 present the quartiles of the distribution of the elasticity with respect to R&D expenditures along with a weighted average computed as $\frac{1}{T}\sum_t\sum_j \mu_{jt}\frac{\partial g(\omega_{jt-1},r_{jt-1})}{\partial r_{jt-1}}$, 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. Excluding industry 2, the returns to R&D at the first, second, and third quartile range between -0.024 and 0.019, -0.003 and 0.024, and 0.015 and 0.051, respectively. Their average is close to 0.012, varying from 0.003 to 0.025 across industries.

Note that negative returns to R&D are legitimate and meaningful in our setting, although some of them may be an artifact of the nonparametric 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.

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 nonperformers using $g_0(\cdot)$ that depends solely on already attained productivity. Columns (5)–(10) of Table 8 summarize the distributions for performers and nonperformers.

Again there is a considerable amount of variation across industries and the firms within an industry. Nevertheless, nonperformers enjoy a systematically higher degree of persistence than performers in industries 1, 2, 3, 4, and 7. An intuitive explanation for this finding is that nonperformers 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.

To facilitate the comparison with the existing literature, we have estimated the knowledge capital model as given in equation (4).[20] Column (11) of Table 8 presents the estimate

[20] We specify a different constant for performers and nonperformers as well as a different time trend (industries 2, 3, 6, and 10) or set of dummies (industries 1, 7, and 8). We drop the term εc_{jt} from equation (4) for nonperformers. To facilitate estimation we impose the widely accepted constraint of constant returns

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 (4) 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 capital tends to be small and rarely significant in the gross-output version but becomes larger in the value-added version. The estimates turn out to be on the low side for this type of exercise. One possible reason may be the non self-selected character of the sample, but perhaps this is the magnitude of estimates that one should expect given the low R&D intensity of Spanish manufacturing firms. Beneito (2001) and Ornaghi (2006), for example, estimate aggregate elasticities ranging from 0.04 to 0.10.

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 8 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 (see also Pakes & Schankerman 1984*b*). Moreover, we find that there are substantial differences between firms in the degree of persistence.

The degree of persistence is expected to be lower when process innovations are rapidly spread or when product innovations are quickly imitated or superseded. (Since output is measured in dollars, we are unable to distinguish between product and process innovations, similar to the knowledge capital literature.) On the other hand, the demand advantage of a product innovation may be offset by a productivity disadvantage if newer products are costlier to produce, thereby lessening the impact of product innovations on persistence.[21] The heterogeneity across firms and industries in the degree of persistence points to an interesting avenue for future research that explores the link between the dynamics of productivity and the nature of product market competition.

One could also argue that the lower degree of persistence is a result of the substantial

to scale in the conventional inputs.

[21]The ESEE survey asks firms whether they have introduced a new product or process over the course of the survey year. This data suggests that, at the level of the industry, the degree of persistence is negatively related to the prevalence of both product and process innovations.

variability in the R&D expenditures that drive the evolution of productivity. The knowledge capital model constructs the stock of knowledge capital that is much smoother and less variable than R&D expenditures. Our view is that the variability in R&D expenditures across firms and periods is likely to contain useful information on the impact of R&D on productivity, but we acknowledge that some of the variability in the R&D expenditures is an artifact of accounting conventions.

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.

Productivity dynamics and industrial change. To explore the dynamics of productivity in more detail, we compute the year-to-year transitions in productivities as follows: We first place firms into five bins according to their actual productivity or flag them as entrants or exitors. We define the bins by the quintiles of the productivity distribution and compute them separately for each year in the sample period. The transition matrix then gives the shares of incumbents that move between the five productivity bins or exit the industry. In addition we record the shares of entrants in the five productivity bins.

Table 9 summarizes the degree of persistence and mobility in the dynamics of productivity. Columns (1) and (2) give the probability of remaining at the bottom and top, respectively, of the productivity distribution along with the expected duration of a stay there. Column (3) gives the probability that a randomly selected firm leaves its productivity bin in the next year. Columns (4)–(6) are analogous to columns (1)–(3) but use the productivity estimates from the knowledge capital model. As can be seen, the knowledge capital model implies a higher degree of persistence at both tails of the productivity distribution (with the possible exception of industry 8). For example, while a firm can expect to remain at the top of its industry for 3 or 4 years, the knowledge capital model implies expected durations of up to 16 years. The knowledge capital model similarly implies a lower degree of mobility. Overall, productivity dynamics are decidedly more sluggish when viewed through the lens of the knowledge capital model. This is because knowledge accumulates from period to period with certainty in the knowledge capital model. In contrast to this deterministic trajectory, in our model the evolution of productivity is stochastic. While firms exert some degree of control over their productivity through their R&D expenditures, they are repeatedly subjected to shocks that render their fortunes less predictable.

To relate productivity dynamics and industrial change, we predict—however crudely—the evolution of an industry. From the transition matrix we compute the shares of firms in the five productivity bins in the steady state of the industry. We tabulate this predicted distribution of performers in columns (7)–(11) of Table 9 along with the actual distribution in our sample in order to describe the initial conditions. We caution the reader that our predictions are extrapolations from the sample period. This, in particular, presumes that the pattern of entry and exit remains unchanged as does the composition of firms that invest in physical capital and knowledge and the corresponding investment intensities. In what follows we do not consider industries 9 and 10 because of the small number of observations with R&D in these industries with low innovative activity.

Nonperformers are fairly evenly spread across the five productivity bins, both in the sample period and in the long run. In contrast, as can be seen from columns (7)–(11) of Table 9, performers tend to be less concentrated at the bottom and more concentrated at the top of the productivity distribution in the sample period and are likely to remain so in the future. This is in line with the fact that the expected productivity of firms that perform R&D is systematically more favorable than that of firms that do not perform R&D in most industries (see Section 5.4 for details). Comparing the predicted productivity distribution in the steady state with the actual productivity distribution in the sample period suggests that the scope for industrial change is limited in many industries. The exceptions are industries 2, 3, and 6 where in the future the share of performers at the top of the productivity distribution increases. In these industries R&D expenditures drive productivity dynamics that, in turn, drive industrial change. At the same time, in industries 7 and 8 the predicted productivity distribution suggests a smaller role for performers in the future, consistent with the fact that productivity growth is lower or almost equal for performers than for nonperformers in these industries with intermediate innovative activity (see Section 5.5 for details).

It is worth noting that, according to the productivity estimates from the knowledge capital model, the outlook for industrial change is much more optimistic. The predicted productivity distribution suggest a much larger role for performers in the future in industries 1, 2, 3, 4, and 6. Indeed, the share of performers at the top of the productivity distribution is predicted to exceed 50% in all industries but 4 and 8, a prediction that we regard as somewhat implausible.

5.7 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 (6) can be decomposed as

$$\beta_t + g(\omega_{jt-1}, r_{jt-1}) - \omega_{jt-1} = [\beta_t + g(\omega_{jt-1}, r_{jt-1}) - g(\omega_{jt-1}, \underline{r})] + [g(\omega_{jt-1}, \underline{r}) - \omega_{jt-1}], \quad (7)$$

where $\underline{r}$ denotes a negligible amount of R&D expenditures.[22] The first term in brackets reflects the change in expected productivity that is attributable to R&D expenditures r_{jt-1}, the second the change that takes place in the absence of investment in knowledge. That is,

[22]Recall 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 a percentile of the industry's R&D expenditures.

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 (7) 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.[23] 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 & Regev 1995, Griliches 2000) and because value added is closer to profits than gross output. We similarly compute the gross rate of return to R&D and the compensation for depreciation from the growth decomposition in equation (7) by multiplying and dividing through by V_{jt} and R_{jt-1}.

Columns (1)–(3) of Table 10 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 35% and differ across industries, ranging from very modest values near 10% to 50%. 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 (4) of Table 10. 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.[24] Column (5) 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. This again suggests that investment in knowledge is systematically more uncertain than investment in physical capital.

[23] 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.

[24] 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 $\mu_j t = \bar{I}_{jt}/\sum_j \bar{I}_{jt}$ are given by the share of investment in physical capital of a firm. In practice, we use firm-specific averages of value added, the stock of physical capital, and 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 the log of the ratio $\xi_{jt}^2/Var(\omega_{jt})$ 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 (6) of Table 10).[25] Moreover, our estimates suggest that engaging in R&D roughly doubles the degree of uncertainty. In contrast, they show a negative impact of investment in physical capital in industries 2, 8, and 10 and no impact in the remaining industries (see column (7)). Investment in physical capital therefore reduces the degree of uncertainty in the productivity process or leaves it unchanged, whereas investment in knowledge enhances it substantially.

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.

To facilitate the comparison with the existing literature, we have used the value-added version of the knowledge capital model 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.[26] The estimated coefficient of this ratio can be interpreted as the gross rate of return to R&D.[27] We obtain the net rate of return to R&D by subtracting the rate of depreciation of knowledge capital. As can be seen from column (8) of Table 10, while the net rates are imprecisely estimated in the knowledge capital model, at around 75% they tend to be higher than the net rates in our model. The question is then whether and why our rates of return to R&D should be considered more reliable and whether this justifies the extra effort of pursuing the more structural approach.

Our rates are computed from more reliable coefficient estimates than what the knowl-

[25] 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.

[26] We specify a different time trend (industries 2, 3, 6, and 10) or set of dummies (industries 1, 7, and 8) for performers and nonperformers. To facilitate estimation we impose the widely accepted constraint of constant returns to scale in the conventional inputs.

[27] 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 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.

edge capital model provides because our estimator takes into account the possibility of endogeneity bias in assessing the role of R&D. Because our model is structural we are more confident in the causality of the estimated relationship between expected productivity, current productivity, and R&D expenditures. The drawback of our approach is that it depends on the informational and timing assumptions that we make. These assumptions, however, appear to be broadly accepted in the literature following OP.

More generally, 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.

R&D policy. While the knowledge capital model yields an industry-wide average rate of return to R&D, our model allows us to recover the entire distribution of rates. The fact that rates differs across firms makes our model potentially useful in determining the allocation of subsidies, a major issue in R&D policy. To illustrate, we have run a reduced-form regression of the net rate of return on the characteristics of the firm that are observable to a policy maker, including polynomials in the size of the firm and its R&D expenditures, the nature of innovation (process vs. product), the R&D employment of the firm, the proportion of R&D subsidies that it receives, the age of the firm, and its investment in physical capital. This regression indicates a systematic relationship with firm size but not with the other variables. Given that firm size and the net rate of return are positively correlated (industries 2, 3, 6, 8, and 9), we tentatively conclude that, if the goal is to maximize returns, then larger firms should be subsidized more extensively than smaller firms. A fuller exploration of the implications of the heterogeneity across firms for R&D policy is left to future research.

6 Concluding remarks

Our goal in this paper 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 that accounts for investment in knowledge. We also derive an estimator for production functions in this setting. We illustrate our approach on 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. By accounting for uncertainty and nonlinearity, our approach extends the knowledge capital model. In fact, the knowledge capital model in its basic form is a special case of our model, albeit one that is rejected by the data. Our model is richer, in particular with regard to the treatment of heterogeneity, thereby allowing us to show that R&D is a major determinant of the differences in productivity across firms and the evolution of firm-level productivity over time. Productivity appears to be considerably more fluid than what the knowledge capital literature suggests, and we tentatively conclude that the scope for industrial change is limited. Our approach also appears to provide us with more plausible answers to questions regarding the rate of return to R&D. 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 these uncertainties are economically significant and matter for firms' investment decisions.

In concluding we emphasize that our approach has already been useful in other contexts. Our model of endogenous productivity change has been applied by Aw, Roberts & Xu (2009) to examine the impact of export market participation on the productivity of firms; other applications are Maican & Orth (2008) and Añón & Manjón (2009). Our estimator for production functions has been extended by Doraszelski & Jaumandreu (2009) to assess the bias of technological change by measuring, at the level of the individual firm, how much of technological change is factor neutral and how much of it is labor augmenting.

Appendix A

Our data come from the ESEE survey. 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 1879 firms in 9 industries. The number of firms with 2, 3,..., 10 years of data is 260, 377, 246, 192, 171, 135, 128, 155, and 215 respectively. Table A1 gives the industry definitions along with their equivalent definitions in terms of the ESEE, National Accounts, and ISIC classifications (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 firms. Total R&D expenditures 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 nonnegligible 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.

- *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 $K_{jt} = \frac{\widetilde{K}_{jt}}{p_{It}}$.

- *Labor.* 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 working 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 labor 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.

- *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.

- *Market dynamism.* Firms are asked to assess the current and future situation (slump, stability, or expansion) of up to 5 separate markets in which they operate. The market dynamism index is computed as a weighted average of the responses.

References

Añón, D. & Manjón, M. (2009), Internationalization, R&D and productivity, Working paper, University of Valencia, Valencia.

Ackerberg, D., Benkard, L., Berry, S. & Pakes, A. (2007), Econometric tools for analyzing market outcomes, *in* J. Heckman & E. Leamer, eds, 'Handbook of Econometrics', Vol. 6A, North-Holland, Amsterdam, pp. 4171–4276.

Ackerberg, D., Caves, K. & Frazer, G. (2006), Structural identification of production functions, Working paper, UCLA, Los Angeles.

Ai, C. & Chen, X. (2003), 'Efficient estimation of models with conditional moment restrictions containing unknown functions', *Econometrica* **71**(6), 1795–1843.

Aw, B., Roberts, M. & Xu, D. (2009), R&D investment, exporting, and productivity dynamics, Working paper, New York University, New York.

Baily, N., Hulten, C. & Campbell, D. (1992), 'Productivity dynamics in manufacturing plants', *Brookings Papers on Economic Activity: Microeconomics* pp. 187–267.

Barret, G. & Donald, S. (2003), 'Consistent tests for stochastic dominance', *Econometrica* **71**(1), 71–104.

Beneito, P. (2001), 'R&D, productivity and spillovers at the firm level: Evidence from Spanish panel data', *Investigaciones Economicas* **25**(2), 289–313.

Blundell, R. & Bond, S. (2000), 'GMM estimation with persistent panel data: An application to production functions', *Econometric Review* **19**, 321–340.

Bond, S. & 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.

Delgado, M., Farinas, J. & 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. & Shleifer, A. (2002), 'The regulation of entry', *Quarterly Journal of Economics* **117**(1), 1–37.

Dolado, J., Garcia-Serrano, C. & Jimeno, J. (2002), 'Drawing lessons from the boom of temporary jobs in Spain', *Economic Journal* **112**(480), 270–295.

Doraszelski, U. & Jaumandreu, J. (2009), Measuring the bias of technological change, Working paper, Harvard University, Cambridge.

European Commission (2001), 'Statistics on innovation in Europe', Enterprise DG, Brussels.

European Commission (2004), 'European competitiveness report', Enterprise DG, Brussels.

Gonzalez, X., Jaumandreu, J. & Pazo, C. (2005), 'Barriers to innovation and subsidy effectiveness', *Rand Journal of Economics* **36**(4), 930–950.

Greenstreet, D. (2005), Exploiting sequential learning to estimate establishment-level productivity dynamics and decision rules, Working paper, University of Michigan, Ann Arbor.

Griffith, R., Harrison, R. & Van Reenen, J. (2006), 'How special is the special relationship? Using the impact of U.S. R&D spillovers on U.K. firms as a test of technological sourcing', *American Economic Review* **96**(5), 1859–1875.

Griliches, Z. (1979), 'Issues in assessing the contribution of R&D to productivity growth', *Bell Journal of Economics* **10**(1), 92–116.

Griliches, Z. (1995), R&D and productivity, *in* P. Stoneman, ed., 'Handbook of the economics of innovation and technical change', Blackwell, Oxford.

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

Griliches, Z. (2000), *R&D, education, and productivity: A retrospective*, Harvard University Press, Cambridge.

Griliches, Z. & Mairesse, J. (1998), Production functions: The search for identification, *in* S. Strom, ed., 'Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium', Cambridge University Press, Cambridge.

Griliches, Z. & Regev, H. (1995), 'Firm productivity in Isreali industry 1979–1988', *Journal of Econometrics* **65**, 175–203.

Hall, B. & Hayashi, F. (1989), Research and development as an investment, Working paper no. 2973, NBER, Cambridge.

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

Hall, R. (1990), Invariance properties of Solow's productivity residual, *in* P. Diamond, ed., 'Growth, productivity, unemployment', MIT Press, Cambridge.

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

Jaffe, A. (1986), 'Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value', *American Economic Review* **76**(5), 984–1001.

Jaumandreu, J. & Mairesse, J. (2005), Using price and demand information to identify production functions, Working paper, Universidad Carlos III de Madrid, Madrid.

Judd, K. (1998), *Numerical methods in economics*, MIT Press, Cambridge.

Klette, T. (1996), 'R&D, scope economies, and plant performance', *Rand Journal of Economics* **27**(3), 502–522.

Levinsohn, J. & Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *Review of Economic Studies* **70**(2), 317–341.

Maican, F. & Orth, M. (2008), Productivity dynamics and the role of "big-box" entrants in retailing, Working paper, University of Gothenburg, Göteborg.

Mairesse, J. & Sassenou, M. (1991), R&D and productivity: A survey of econometric studies at the firm level, Working paper no. 3666, NBER, Cambridge.

Marschak, J. & Andrews, W. (1944), 'Random simultaneous equations and the theory of production', *Econometrica* **12**(3), 143–205.

Nadiri, M. (1993), Innovations and technical spillovers, Working paper, New York University, New York.

Nadiri, M. & Prucha, I. (1996), 'Estimation of depreciation rate of physical and R&D capital in the U.S. total manufacturing sector', *Economic Inquiry* **34**(1), 45–56.

Newey, W. & Powell, J. (2003), 'Instrumental variables estimation on nonparametric models', *Econometrica* **71**(5), 1565–1578.

Newey, W., Powell, J. & Vella, F. (1999), 'Nonparametric estimation of triangular simltaneous equations models', *Econometrica* **67**(3), 565–603.

OECD (2007), 'The policy mix for research, development and innovation in Spain: Key issues and policy recommendations', Directorate for Science, Technology and Industry, Paris.

Olley, S. & Pakes, A. (1996), 'The dynamics of productivity in the telecommunications industry', *Econometrica* **64**(6), 1263–1297.

Ornaghi, C. (2006), 'Spillovers in product and process innovation: Evidence from manufacturing firms', *International Journal of Industrial Organization* **24**(2), 349–380.

Pakes, A. & 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**(2), 243–262.

Pakes, A. & Schankerman, M. (1984*a*), An exploration into the determinants of research intensity, *in* Z. Griliches, ed., 'R&D, patents and productivity', University of Chicago Press, Chicago.

Pakes, A. & Schankerman, M. (1984*b*), The rate of obsolescence of patents, research gestation lags, and the private rate of return to research resources, *in* Z. Griliches, ed., 'R&D, patents and productivity', University of Chicago Press, Chicago.

Pavcnik, N. (2002), 'Trade liberalization, exit, and productivity improvement: Evidence from Chilean plants', *Review of Economic Studies* **69**(1), 245–276.

Rivers, D. & Vuong, Q. (2002), 'Model selection tests for nonlinear dynamic models', *Econometrics Journal* **5**(1), 1–39.

Robinson, P. (1988), 'Root-n-consistent semiparametric regression', *Econometrica* **56**(4), 931–954.

Topalova, P. (2004), Trade liberalization and firm productivity: The case of India, Working paper no. 04/28, IMF, Washington.

Wooldridge, J. (2004), On estimating firm-level production functions using proxy variables to control for unobservables, Working paper, Michigan State University, East Lansing.

Table 1: Descriptive statistics.

Industry	Obs.[a]	Firms[a]	Entry[a] (%)	Exit[a] (%)	Rates of growth[a] Output (s. d.)	Labor (s. d.)	Capital (s. d.)	Materials (s. d.)	Price (s. d.)	With R&D[b] Obs. (%)	Stable (%)	Occas. (%)	R&D intensity (s. d.)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1. Metals and metal products	1235	289	88	17	0.050	0.010	0.086	0.038	0.012	420	63	72	0.0126
			(30.4)	(5.9)	(0.238)	(0.183)	(0.278)	(0.346)	(0.055)	(34.0)	(21.8)	(24.9)	(0.0144)
2. Non-metallic minerals	670	140	20	15	0.039	0.002	0.065	0.040	0.011	226	22	44	0.0112
			(14.3)	(10.7)	(0.209)	(0.152)	(0.259)	(0.304)	(0.057)	(33.7)	(15.7)	(31.4)	(0.0206)
3. Chemical products	1218	275	64	15	0.068	0.007	0.093	0.054	0.007	672	124	55	0.0268
			(23.3)	(5.5)	(0.196)	(0.146)	(0.238)	(0.254)	(0.061)	(55.2)	(45.1)	(20.0)	(0.0353)
4. Agric. and ind. machinery	576	132	36	6	0.059	0.010	0.078	0.046	0.013	322	52	35	0.0219
			(27.3)	(4.5)	(0.275)	(0.170)	(0.247)	(0.371)	(0.032)	(55.9)	(39.4)	(26.5)	(0.0275)
6. Transport equipment	637	148	39	10	0.087	0.011	0.114	0.087	0.007	361	62	35	0.0224
			(26.4)	(6.8)	(0.0354)	(0.207)	(0.255)	(0.431)	(0.037)	(56.7)	(41.9)	(23.6)	(0.0345)
7. Food, drink and tobacco	1408	304	47	22	0.025	-0.003	0.094	0.019	0.022	386	56	64	0.0071
			(15.5)	(7.2)	(0.224)	(0.186)	(0.271)	(0.305)	(0.065)	(27.4)	(18.4)	(21.1)	(0.0281)
8. Textile, leather and shoes	1278	293	77	49	0.020	-0.007	0.059	0.012	0.016	378	39	66	0.0152
			(26.3)	(16.7)	(0.233)	(0.192)	(0.235)	(0.356)	(0.040)	(29.6)	(13.3)	(22.5)	(0.0219)
9. Timber and furniture	569	138	52	18	0.038	0.014	0.077	0.029	0.020	66	7	18	0.0138
			(37.7)	(13.0)	(0.278)	(0.210)	(0.257)	(0.379)	(0.035)	(12.6)	(5.1)	(13.8)	(0.0326)
10. Paper and printing products	665	160	42	10	0.035	-0.000	0.099	0.026	0.019	113	21	25	0.0143
			(26.3)	(6.3)	(0.183)	(0.140)	(0.303)	(0.265)	(0.089)	(17.0)	(13.1)	(13.8)	(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.

Table 2: Production function estimates.

Industry	OLS[a]			Exogenous Markov process[a]			Controlled Markov process[a]		
	Labor (std. err.)	Capital (std. err.)	Materials (std. err.)	Labor (std. err.)	Capital (std. err.)	Materials (std. err.)	Labor (std. err.)	Capital (std. err.)	Materials (std. err.)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Metals and metal products	0.252	0.109	0.642	0.171	0.110	0.677	0.097	0.122	0.684
	(0.022)	(0.013)	(0.020)	(0.034)	(0.020)	(0.015)	(0.021)	(0.011)	(0.008)
2. Non-metallic minerals	0.277	0.091	0.662	0.166	0.148	0.655	0.147	0.204	0.633
	(0.032)	(0.020)	(0.028)	(0.033)	(0.021)	(0.015)	(0.016)	(0.015)	(0.015)
3. Chemical products	0.239	0.060	0.730	0.149	0.123	0.723	0.113	0.131	0.719
	(0.021)	(0.010)	(0.020)	(0.032)	(0.021)	(0.013)	(0.021)	(0.014)	(0.009)
4. Agric. and ind. machinery	0.284	0.051	0.671	0.261	0.073	0.636	0.276	0.077	0.641
	(0.038)	(0.017)	(0.027)	(0.041)	(0.019)	(0.016)	(0.029)	(0.015)	(0.013)
6. Transport equipment	0.289	0.080	0.636	0.194	0.101	0.679	0.147	0.135	0.657
	(0.033)	(0.023)	(0.046)	(0.037)	(0.023)	(0.016)	(0.010)	(0.007)	(0.008)
7. Food, drink and tobacco	0.177	0.094	0.739	0.175	0.046	0.764	0.116	0.081	0.760
	(0.016)	(0.014)	(0.016)	(0.021)	(0.014)	(0.009)	(0.016)	(0.011)	(0.007)
8. Textile, leather and shoes	0.335	0.059	0.605	0.265	0.056	0.611	0.271	0.059	0.605
	(0.024)	(0.011)	(0.019)	(0.037)	(0.020)	(0.015)	(0.023)	(0.012)	(0.010)
9. Timber and furniture	0.283	0.079	0.670	0.174	0.121	0.705	0.176	0.131	0.697
	(0.029)	(0.019)	(0.029)	(0.021)	(0.014)	(0.017)	(0.017)	(0.009)	(0.011)
10. Paper and printing products	0.321	0.092	0.621	0.265	0.142	0.593	0.249	0.121	0.617
	(0.029)	(0.016)	(0.025)	(0.027)	(0.016)	(0.018)	(0.025)	(0.013)	(0.014)

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

Table 3: Specification tests.

Industry	Perfect competition test			Overidentifying restrictions: All		Lagged labor		Lagged materials		Capital and lagged capital		Parametric inversion test	
	$\chi^2(9)$	p val.	η (std. dev.)	$\chi^2(df)$	p val.	$\chi^2(df)$	p val.	$\chi^2(1)$	p val.	$\chi^2(df)$	p val.	$\chi^2(3)$	p val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1. Metals and metal products	238.05	0.000	2.09	109.24	0.039	28.14	0.708	0.00	0.980	49.51	0.042	5.19	0.158
			(0.12)	(85)		(33)				(34)			
2. Non-metallic minerals	127.19	0.000	1.75	52.10	0.137	13.39	0.768	1.21	0.271	35.97	0.011	3.74	0.291
			(0.15)	(42)		(18)				(19)			
3. Chemical products	23.80	0.005	1.99	95.22	0.210	30.49	0.593	1.45	0.229	33.77	0.479	0.92	0.821
			(0.03)	(85)		(33)				(34)			
4. Agric. and ind. machinery	85.57	0.000	1.90	45.47	0.369	26.31	0.093	0.15	0.699	19.01	0.456	2.04	0.564
			(0.11)	(43)		(18)				(19)			
6. Transport equipment	762.66	0.000	1.93	88.79	0.368	36.84	0.296	0.04	0.841	35.87	0.381	7.08	0.069
			(0.10)	(85)		(33)				(34)			
7. Food, drink and tobacco	29.85	0.000	2.12	97.57	0.166	33.20	0.458	0.66	0.417	39.67	0.232	7.56	0.056
			(0.09)	(85)		(33)				(34)			
8. Textile, leather and shoes	144.10	0.000	1.91	96.76	0.180	29.28	0.653	1.28	0.258	35.93	0.378	0.62	0.892
			(0.09)	(85)		(33)				(34)			
9. Timber and furniture	102.80	0.000	1.81	44.95	0.390	18.50	0.423	0.00	0.997	27.08	0.103	0.62	0.892
			(0.10)	(43)		(18)				(19)			
10. Paper and printing products	61.57	0.000	1.95	51.37	0.152	22.24	0.222	0.75	0.386	34.98	0.014	5.88	0.118
			(0.15)	(42)		(18)				(19)			

Table 4: Nonlinearity and uncertainty.

Industry	Exogeneity test		Separability test		Complements/ substitutes[a]		Scale economies[a]			
	$\chi^2(10)$	p val.	$\chi^2(3)$	p val.	%obs.>0	%obs.<0	%obs.>0	%obs.<0	$\frac{Var(e)}{Var(\omega)}$	$\frac{Var(\xi)}{Var(\omega)}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Metals and metal products	181.21	0.000	18.70	0.000	0.52	0.08	0.08	0.01	0.772	0.589
2. Non-metallic minerals	56.93	0.000	6.99	0.072	0.19	0.09	0.00	0.00	0.916	0.557
3. Chemical products	98.88	0.000	19.92	0.000	0.44	0.10	0.21	0.47	0.725	0.221
4. Agric. and ind. machinery	53.39	0.000	11.20	0.011	0.29	0.13	0.02	0.21	1.390	0.494
6. Transport equipment	780.71	0.000	281.75	0.000	0.49	0.32	0.64	0.21	1.338	0.544
7. Food, drink and tobacco	111.90	0.000	26.40	0.000	0.12	0.27	0.52	0.15	1.369	0.265
8. Textile, leather and shoes	104.55	0.000	16.79	0.001	0.03	0.21	0.72	0.02	1.148	0.308
9. Timber and furniture	118.20	0.000	39.82	0.000	0.44	0.15	0.35	0.23	1.417	0.515
10. Paper and printing products	59.73	0.000	106.92	0.000	0.64	0.05	0.45	0.06	0.713	0.433

[a] We compute the variance of the derivative using the δ-method.

Table 5: Knowledge capital model tests.

Industry	Basic model		Generalized model 1		Generalized model 2	
	$N(0,1)$	p val.	$N(0,1)$	p val.	$N(0,1)$	p val.
	(1)	(2)	(3)	(4)	(5)	(6)
1. Metals and metal products	...	0.000	...	0.000	-17.77	0.000
2. Non-metallic minerals	...	0.000	...	0.000	-9.15	0.000
3. Chemical products	...	0.000	...	0.000	-14.91	0.000
4. Agric. and ind. machinery	...	0.000	...	0.000	-10.94	0.000
6. Transport equipment	...	0.000	...	0.000	-8.65	0.000
7. Food, drink and tobacco	...	0.000	...	0.000	-8.10	0.000
8. Textile, leather and shoes	...	0.000	...	0.000	-17.65	0.000
9. Timber and furniture	...	0.000	...	0.000	-11.26	0.000
10. Paper and printing products	...	0.000	...	0.000	-18.72	0.000

Table 6: Productivity levels.

Industry	Size	Firms with		Diff. of means	Standard deviation		Mean with R&D is greater		Var. with R&D is greater		Kolmogorov-Smirnov test: Distributions are equal		Kolmogorov-Smirnov test: Distribution with R&D dominates	
		No R&D	R&D		No R&D	R&D	t	p val.	F	p val.	S_1	p val.	S_2	p val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Metals and metal products	≤ 200	143	71	0.044	0.102	0.095	-5.193	1.000	1.147	0.147	2.066	0.000	0.327	0.807
	> 200	11	64	0.026	0.123	0.109	-1.970	0.974	1.282	0.052				
2. Non-metallic minerals	≤ 200	65	27	0.055	0.174	0.173	-2.593	0.994	1.016	0.478	1.341	0.055	0.697	0.379
	> 200	9	39	0.002	0.139	0.120	-0.073	0.529	1.338	0.051				
3. Chemical products	≤ 200	91	81	0.040	0.098	0.081	-5.814	1.000	1.471	0.000	1.350	0.052	0.242	0.889
	> 200	5	98	0.033	0.097	0.098	-2.305	0.987	0.981	0.516				
4. Agric. and ind. machinery	≤ 200	39	56	-0.016	0.113	0.134	1.253	0.106	0.715	0.990	0.935	0.346	0.935	0.174
	> 200	6	31	-0.067	0.112	0.089	3.174	0.002	1.601	0.032				
6. Transport equipment	≤ 200	37	32	0.069	0.098	0.078	-6.183	1.000	1.551	0.011	1.564	0.015	0.000	1.000
	> 200	14	65	0.047	0.088	0.075	-4.893	1.000	1.372	0.022				
7. Food, drink and tobacco	≤ 200	155	49	0.039	0.116	0.098	-3.715	1.000	1.400	0.018	2.062	0.000	0.315	0.820
	> 200	29	71	0.039	0.092	0.082	-4.874	1.000	1.245	0.046	0.820	0.512	0.123	0.970
8. Textile, leather and shoes	≤ 200	165	59	0.035	0.094	0.123	-3.639	1.000	0.590	1.000	1.643	0.009	0.268	0.866
	> 200	23	46	0.004	0.117	0.111	-0.390	0.651	1.125	0.211	0.681	0.743	0.596	0.492
9. Timber and furniture	≤ 200	112	18	-0.020	0.098	0.153	0.724	0.237	0.407	1.000				
	> 200	1	7	0.014	0.025	0.121	-0.611	0.721	0.043	1.000				
10. Paper and printing products	≤ 200	98	24	-0.012	0.092	0.148	0.590	0.279	0.383	1.000	0.930	0.353	0.709	0.366
	> 200	16	22	0.028	0.100	0.137	-1.400	0.917	0.530	0.998				

Table 7: Productivity growth.

Industry	Unweighted productivity growth[a]			Weighted productivity growth[a]		
	Total (std. dev.)	R&D obs. (std. dev.)	No R&D obs. (std. dev.)	Total	R&D obs. (%contrib.)	No R&D obs. (%contrib.)
	(1)	(2)	(3)	(4)	(5)	(6)
1. Metals and metal products	0.0140	0.0174	0.0131	0.0141	0.0171	0.0099
	(0.0459)	(0.0423)	(0.0468)		(68.0)	(32.0)
2. Non-metallic minerals	0.0107	0.0229	0.0070	0.0033	0.0054	-0.0002
	(0.0881)	(0.0644)	(0.0938)		(85.1)	(14.9)
3. Chemical products	0.0163	0.0197	0.0144	0.0176	0.0214	0.0098
	(0.0383)	(0.0368)	(0.0390)		(82.9)	(17.1)
4. Agric. and ind. machinery	0.0145	0.0126	0.0161	0.0176	0.0159	0.0236
	(0.0590)	(0.0575)	(0.0602)		(65.8)	(34.2)
6. Transport equipment	0.0237	0.0325	0.0179	0.0276	0.0308	0.0182
	(0.0452)	(0.0474)	(0.0427)		(82.2)	(17.8)
7. Food, drink and tobacco	0.0101	0.0108	0.0100	0.0022	0.0026	0.0017
	(0.0465)	(0.0489)	(0.0462)		(69.3)	(30.7)
8. Textile, leather and shoes	0.0140	-0.0002	0.0179	0.0084	0.0016	0.0129
	(0.0501)	(0.0559)	(0.0476)		(4.4)	(95.6)
9. Timber and furniture	0.0099	0.0368	0.0082	0.0102	0.0386	0.0062
	(0.0565)	(0.0815)	(0.0541)		(46.4)	(53.6)
10. Paper and printing products	0.0123	0.0003	0.0135	0.0089	0.0015	0.0110
	(0.0565)	(0.0797)	(0.0534)		(11.4)	(88.6)

[a] We trim 2.5% of observations at each tail of the distribution.

Table 8: 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} Performers[b]			Elasticity wrt. ω_{jt-1} Nonperformers[b]			Knowledge capital model Elasticity wrt. C_{jt} and R_{jt-1}: Gross output		Knowledge capital model Elasticity wrt. C_{jt} and R_{jt-1}: Value added	
	Q_1	Q_2	Q_3	Mean	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3	ε (s. e.)[c]	Mean	ε (s. e.)[c]	Mean
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Metals and metal products	0.019	0.024	0.029	0.023	0.374	0.474	0.529	0.485	0.743	0.896	0.003 (0.005)	0.001	0.025 (0.016)	0.005
2. Non-metallic minerals	-0.022	-0.019	-0.002	-0.009	0.190	0.265	0.438	0.371	0.683	0.876	0.013 (0.006)	0.002	0.046 (0.016)	0.009
3. Chemical products	0.010	0.013	0.015	0.012	0.553	0.585	0.629	0.589	0.772	0.843	0.018 (0.004)	0.003	0.075 (0.011)	0.014
4. Agric. and ind. machinery	-0.012	-0.003	0.023	0.007	0.409	0.665	0.747	0.681	0.851	0.967	-0.003 (0.008)	-0.001	0.025 (0.016)	0.005
6. Transport equipment	-0.023	0.000	0.016	0.019	0.497	0.622	0.691	0.481	0.568	0.660	0.009 (0.005)	0.002	0.017 (0.017)	0.004
7. Food, drink and tobacco	-0.001	0.013	0.025	0.015	0.598	0.765	0.889	0.784	0.854	0.878	0.002 (0.006)	0.000	0.046 (0.012)	0.011
8. Textile, leather and shoes	0.008	0.018	0.032	0.025	0.694	0.718	0.760	0.553	0.641	0.705	0.011 (0.008)	0.002	0.018 (0.018)	0.003
9. Timber and furniture	-0.022	0.006	0.051	0.009	0.458	0.585	0.814	0.303	0.430	0.641	0.018 (0.011)	0.004	0.074 (0.030)	0.016
10. Paper and printing products	-0.024	0.023	0.049	0.003	0.405	0.676	0.812	0.569	0.644	0.670	0.013 (0.009)	0.003	0.041 (0.028)	0.010

[a] We trim 7.5% of observations at the left tail of the distribution and 2.5% of observations at the right tail.
[b] We trim observations below zero and above unity.
[c] All standard errors are robust to heteroskedasticity and autocorrelation.

Table 9: Productivity dynamics and industrial change.

				Knowledge capital model			With R&D				
	Persistence			Persistence			Productivity distribution				
Industry	At bottom prob. (yrs.)	At top prob. (yrs.)	Mobility prob.	At bottom prob. (yrs.)	At top prob. (yrs.)	Mobility prob.	Bottom	...	Middle	...	Top
							predicted and actual probability				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. Metals and metal products	0.68	0.75	0.41	0.69	0.82	0.29	0.12	0.20	0.13	0.19	0.36
	(3.14)	(4.24)		(3.37)	(5.72)		0.12	0.19	0.13	0.21	0.35
2. Non-metallic minerals	0.71	0.69	0.45	0.79	0.87	0.20	0.14	0.19	0.19	0.19	0.29
	(3.61)	(3.41)		(5.42)	(7.71)		0.20	0.20	0.24	0.17	0.18
3. Chemical products	0.68	0.74	0.42	0.82	0.91	0.22	0.08	0.15	0.17	0.27	0.33
	(3.19)	(4.10)		(6.10)	(12.87)		0.10	0.18	0.20	0.27	0.26
4. Agric. and ind. machinery	0.62	0.57	0.52	0.89	0.87	0.15	0.14	0.19	0.25	0.23	0.19
	(2.67)	(2.48)		(12.34)	(7.80)		0.15	0.18	0.22	0.25	0.20
6. Transport equipment	0.61	0.66	0.50	0.83	0.92	0.22	0.08	0.18	0.18	0.25	0.31
	(2.94)	(3.11)		(6.23)	(13.27)		0.13	0.21	0.18	0.20	0.28
7. Food, drink and tobacco	0.69	0.69	0.42	0.73	0.90	0.22	0.12	0.15	0.25	0.23	0.25
	(3.29)	(3.58)		(4.00)	(11.33)		0.14	0.13	0.20	0.22	0.32
8. Textile, leather and shoes	0.61	0.72	0.44	0.77	0.63	0.40	0.08	0.13	0.22	0.22	0.35
	(2.96)	(4.07)		(4.91)	(2.79)		0.06	0.12	0.19	0.22	0.41
9. Timber and furniture	0.68	0.56	0.50	0.86	0.91	0.09					
	(3.21)	(2.57)		(9.92)	(15.62)						
10. Paper and printing products	0.69	0.69	0.47	0.86	0.86	0.18					
	(3.31)	(3.61)		(7.47)	(8.13)						

Table 10: Rates of return to R&D and investment in physical capital and degree of uncertainty.

Industry	R&D[a]			Physical capital		Regression of $\frac{\xi_{jt}^2}{Var(\omega_{jt})}$ on[b]		Knowledge capital model
	Gross rate	Net rate	Depreciation	net rate	Ratio	R&D (std. err.)[d]	Investment (std. err.)[d]	net rate[c] (std. err.)[d]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Metals and metal products	0.688	0.549	0.139	0.227	2.4	0.698 (0.185)	-0.202 (0.198)	2.025 (1.435)
2. Non-metallic minerals	0.863	0.468	0.395	0.306	1.5	1.170 (0.216)	-0.459 (0.261)	0.684 (0.362)
3. Chemical products	0.924	0.385	0.539	0.198	1.9	0.573 (0.136)	0.027 (0.262)	0.540 (0.249)
4. Agric. and ind. machinery	0.636	0.333	0.303	0.180	1.9	0.831 (0.189)	-0.134 (0.228)	0.685 (0.051)
6. Transport equipment	1.066	0.401	0.665	0.261	1.5	1.368 (0.200)	-0.034 (0.413)	0.558 (0.211)
7. Food, drink and tobacco	1.163	0.102	1.061	0.035	2.9	0.964 (0.169)	-0.043 (0.166)	0.705 (0.184)
8. Textile, leather and shoes	0.418	0.066	0.352	0.043	1.5	1.233 (0.155)	-0.264 (0.152)	0.648 (0.501)
9. Timber and furniture	0.785	0.386	0.399	0.253	1.5	2.072 (0.382)	-0.236 (0.212)	0.099 (0.628)
10. Paper and printing products	0.783	0.472	0.311	0.043	10.9	1.284 (0.274)	-0.455 (0.233)	1.008 (0.098)

[a] We calculate the first two terms of the decomposition in equation (7) and infer the third term. We trim the distribution of first term to retain between 70% and 95% of observations.

[b] We include 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 in the regression.

[c] We trim 2.5% of observations at each tail of the distribution.

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

Table A1: Industry definitions, equivalent classifications, and shares.

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

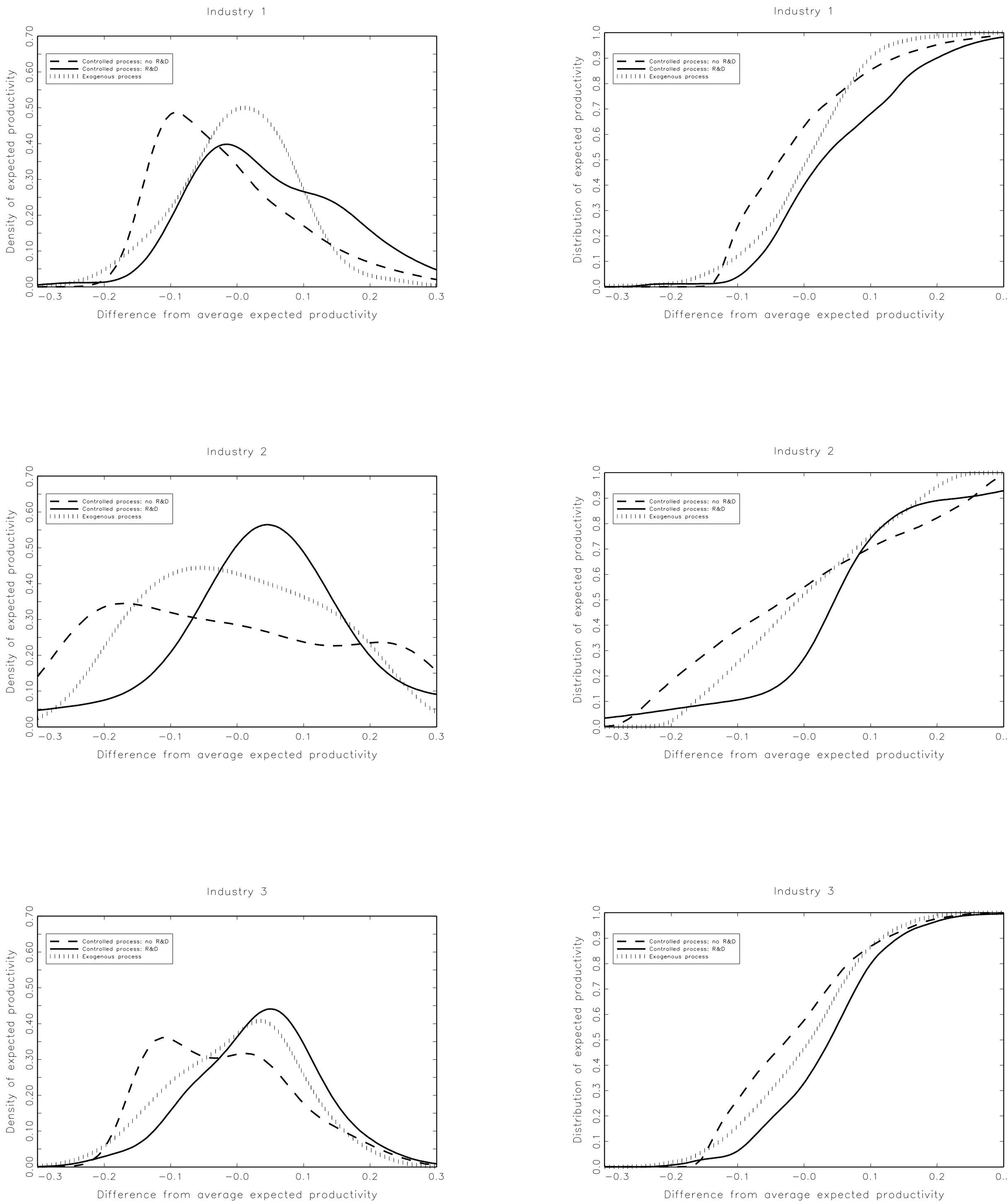

Figure 1: Productivity levels. Density (left panels) and distribution (right panels) of expected productivity.

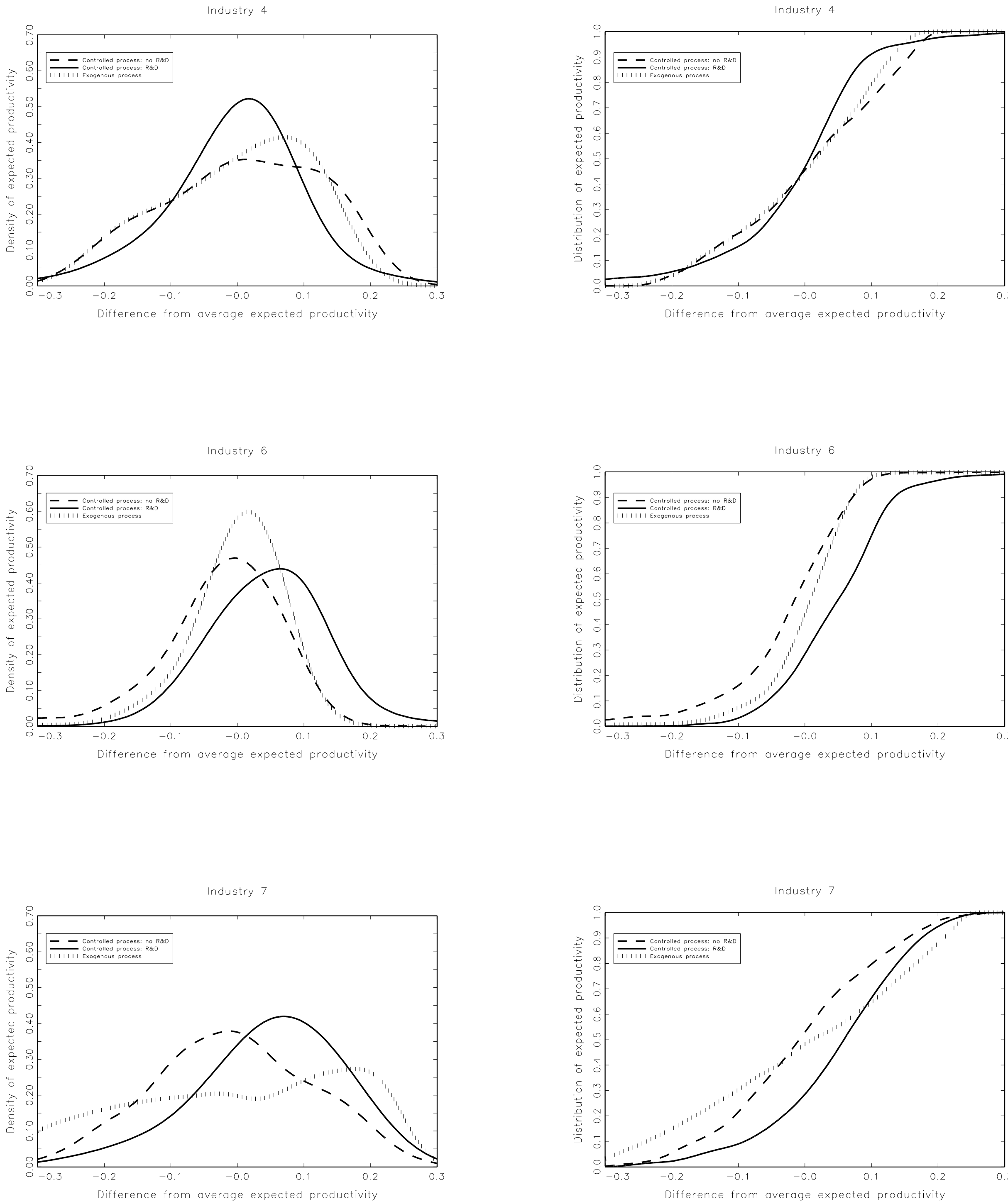

Figure 1: (cont'd) Productivity levels. Density (left panels) and distribution (right panels) of expected productivity.

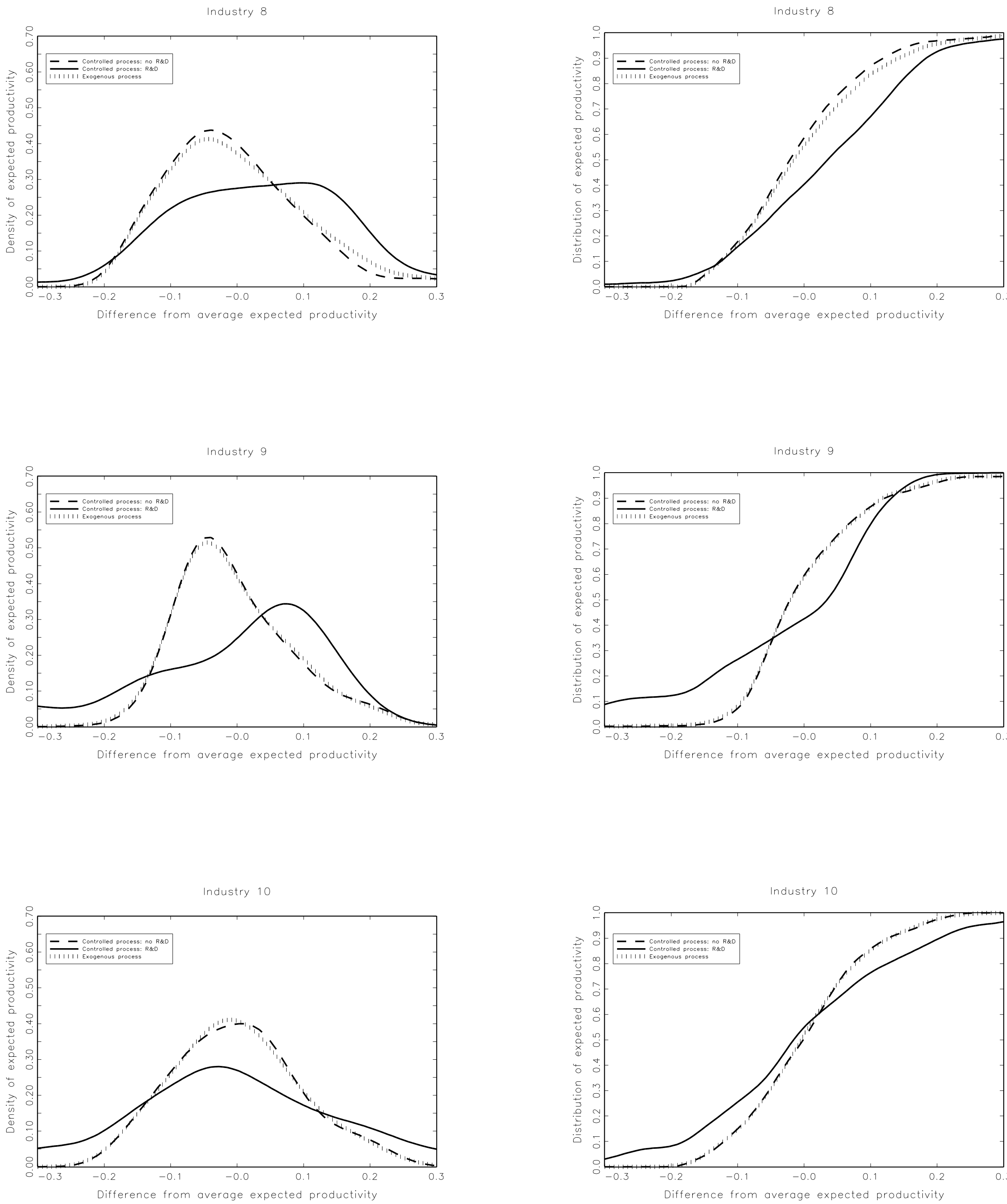

Figure 1: (cont'd) Productivity levels. Density (left panels) and distribution (right panels) of expected productivity.

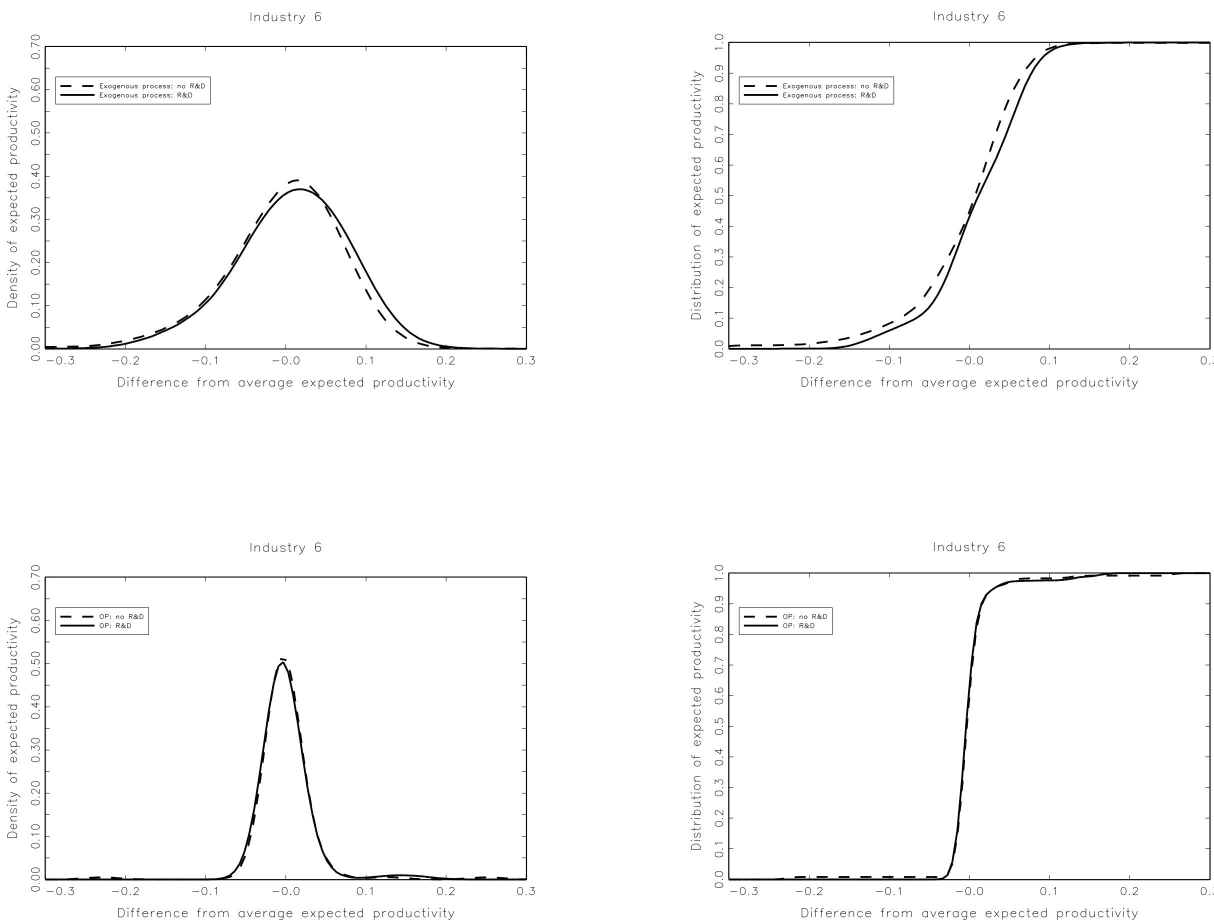

Figure 2: Productivity levels. Density (left panels) and distribution (right panels) of expected productivity. Exogenous Markov process (upper panels) and OP estimator (lower panels).

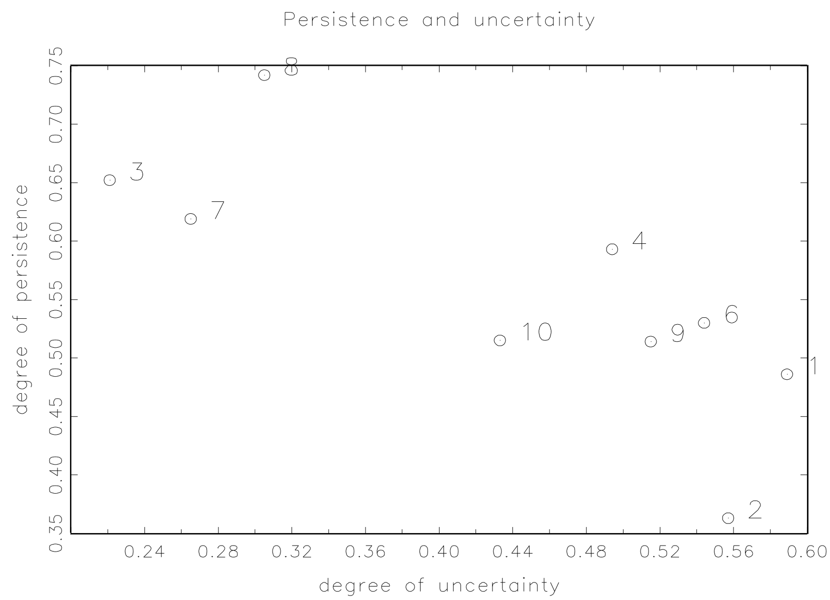

Figure 3: Persistence and uncertainty.

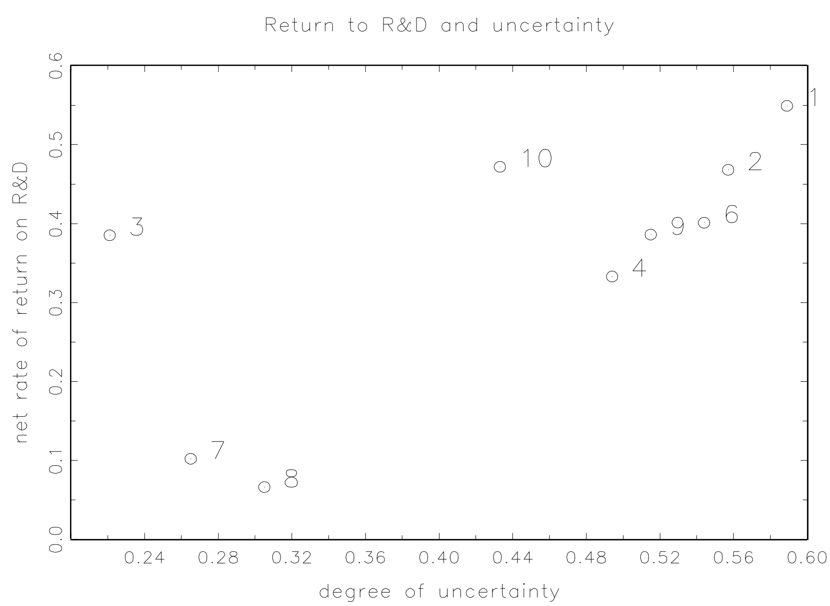

Figure 4: Return to R&D and uncertainty.