Abstract

Investment in knowledge at the firm level is a primary source of productivity growth. Griliches (1979) developed the knowledge capital model to account for the fact that investment in knowledge creates long-lived assets for firms. While the knowledge capital model has remained a cornerstone of the productivity literature for more than 25 years, the links between R&D and productivity are much more complex. In this paper, we revisit the question of the impact of the investment in knowledge on the productivity of firms by relaxing drastically the assumptions on the R&D process. In particular, we recognize the uncertainties in the R&D process in the form of shocks to productivity that accumulate over time.

We develop simple estimators for production functions in the presence of endogenous productivity change. The basic idea is to exploit the fact that decisions on variable inputs such as labor and materials are based on current productivity. This results in input demands that are invertible functions and that can thus be used to control for unobserved productivity in the estimation. Moreover, the parametric specification of the production function implies a known form for these functions. This contrasts with the nonparametric methods in the existing literature and renders identification and estimation more tractable. Perhaps the most apparent benefit of our approach is that it allows us to retrieve productivity and its relationship with R&D at the firm level.

We illustrate our approach to production function estimation on an unbalanced panel of Spanish manufacturing firms that cover a total of more than 1800 firms in nine industries during the 1990s. A number of interesting findings emerge. R&D is subject to a high degree of uncertainty but expected productivity is nevertheless increasing in R&D expenditures and attained productivity. The distribution of expected productivity for firms that perform R&D stochastically dominates the distribution for nonperformers and the growth in expected productivity is higher for performers. Assuming an exogenous process for productivity, as is often done in the existing literature, blurs these and other fundamental differences among firms.

*We thank Dan Ackerberg, David Greenstreet, and Jacques Mairesse for helpful discussions and Laia Castany for excellent research assistance.
†Department of Economics, Harvard University, Littauer Center, 1875 Cambridge Street, Cambridge, MA 02138, USA. E-mail: doraszelski@harvard.edu.
‡Dpto. Economía, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), Spain. E-mail: jordij@eco.uc3m.es.
1 Introduction

Investment in knowledge at the firm level is a primary source of productivity growth. Firms invest in R&D and related activities to develop and introduce process and product innovations that enhance their productivity. The links between R&D and productivity are, however, far from simple. The outcome of the R&D process is likely to be subject to a high degree of uncertainty. Discovery is, by its very nature, uncertain. Once discovered an idea has to be developed and applied, and there are the technical and commercial uncertainties linked to its practical implementation. In addition, current and past investments in knowledge are likely to interact with each other in many ways.

To assess the impact of the investment in knowledge on the productivity of firms, Griliches (1979) developed the knowledge capital model that has remained a cornerstone of the productivity literature for more than 25 years. The main idea is to augment the production function with the stock of knowledge, proxied for in one way or another. The knowledge capital model has been used in various forms in hundreds of empirical studies on firm-level productivity and also applied to macroeconomic growth models (see Griliches (1995) for an extensive survey).\(^1\) While useful as a practical tool, the knowledge capital model has a long list of known drawbacks as explained, for example, in Griliches (2000). The critical (but implicit) assumptions of the basic model include the linear and certain accumulation of knowledge from period to period in proportion to R&D expenditures as well as the linear and certain depreciation.\(^2\)

In this paper, we revisit the question of the impact of the investment in knowledge on the productivity of firms by relaxing drastically the assumptions on the R&D process. In particular, we recognize the uncertainties in the R&D process in the form of shocks to productivity that accumulate over time. We flexibly model the interactions between current and past investments in knowledge and we relax the assumption that the obsolescence of previously acquired knowledge can be described by a continuous function of time.

Our starting point is a dynamic investment model. A firm can invest 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. The evolution of productivity is subject to random shocks. We interpret these innovations to productivity as representing the resolution over time of all uncertainties. They capture the usual factors that have a persistent influence on productivity (e.g., absorption of techniques, modification of processes, gains and losses due to changes in labor composition and management abilities). R&D governs the evolution of productivity up to

---


\(^2\)Subsequent work by Klette (1996) and others has replaced the linear accumulation rule by nonlinear albeit parametric rules.
an unpredictable component. Hence, 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 and success in implementation. Productivity thus follows a first-order Markov process that can be shifted by R&D expenditures. Subsequently decisions on variable (or “static”) inputs such as labor and materials are taken according to the current productivity and capital stock of the firm.

To estimate the parameters of the production function and retrieve productivity at the level of the firm, we build on the recent literature that has explicitly modeled the evolution of productivity over time. Starting with Olley & Pakes (1996) this literature has stressed that (observed) investment depend on (unobserved) productivity.\(^3\) Inverting this relationship allows controlling for productivity, thereby resolving the endogeneity problem in production function estimation as well as, eventually, the selection problem arising from firm exit. In addition to Olley & Pakes (1996) (hereafter OP), this line of research includes contributions by Levinsohn & Petrin (2003) (hereafter LP), Wooldridge (2004), Ackerberg, Caves & Frazer (2005) as well as a long list of applications.

We develop simple estimators for production functions that extend this line of research. Unlike existing estimators we are able to accommodate the controlled Markov process that results from the impact of R&D on the evolution of productivity. The basic idea is to exploit the fact that decisions on variable inputs are based on current productivity. Since these inputs are chosen with current productivity known, they contain information about it. 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. Moreover, given a parametric specification of the production function, the functional form of these inverse input demand functions is known. This contrast with the nonparametric methods in the existing literature and renders identification and estimation more tractable. Perhaps the most apparent benefit of our approach is that it allows us to retrieve productivity and its relationship with R&D at the firm level.

The demand for labor in a value added production function and the demands for both labor and materials in a gross output production function are argued to be adequate controls for unobserved productivity. A way to combine multiple input demands in order to increase the efficiency of the estimator is devised. In addition, we discuss how the specification must be modified when firms are producers of differentiated products in an imperfectly competitive environment.

We illustrate our estimator on an unbalanced panel of Spanish manufacturing firms that cover a total of more than 1800 firms in nine industries during the 1990s. We compare the productivities of firms with different amounts of R&D expenditures. We also quantify the role played by innovative activities in different industries, assess the importance of uncertainty in the R&D process, characterize the growth of productivity over time, and

---

\(^3\)See Griliches & Mairesse (1998) for a review of the problems involved in the estimation of production functions.
compute rates of return to R&D.

A number of interesting findings emerge. To begin with the R&D process must be treated as inherently uncertain. We estimate that, depending on the industry, between 20% and 50% of the variance in actual productivity is explained by productivity innovations that cannot be predicted when R&D expenditures are carried out. Despite this, the distribution of expected productivity of firms that perform R&D tends to stochastically dominate the distribution of expected productivity of firms that do not perform R&D. The mean of the distribution of expected productivity is higher for performers than for nonperformers by around 5% in most cases and up to 9% in some cases. However, R&D may not inject additional uncertainty into the evolution of productivity over time.

Expected productivity is increasing in attained productivity and R&D expenditures. While this relationship takes a simple separable form in some cases, in most cases the impact of current R&D on future productivity depends crucially on current productivity. Hence, current and past investments in knowledge interact in a complex fashion. We also estimate that the contribution of firms that perform R&D explains between 50% and 85% of productivity growth in the industries with intermediate or high innovative activity. R&D expenditures are therefore an important source of productivity growth. Finally, we find large average rates of return to R&D in most industries, although our estimates are within the range of the previous literature. 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 and heterogeneity across firms. Assuming an exogenous process for productivity, as is often done in the existing literature, overlooks some of its most interesting features.

2 A model for investment in knowledge

Many firms carry out two types of investments, one in physical capital and another in knowledge through R&D expenditures. Assume that a firm makes its investment decisions 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

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

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 fixed (or “dynamic”) input among the conventional factors of production, and accumulates according to $K_{jt} = (1-\delta)K_{jt-1} + I_{jt-1}$. This 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 $\omega_{jt}$. We follow OP and often refer to $\omega_{jt}$ 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 $t$.

The assumption usually made about productivity (see OP, LP, and the subsequent literature) is that it follows an exogenous first-order Markov process with transition probabilities $P(\omega_{jt}|\omega_{jt-1})$. This rules out innovative activities. Investment in knowledge has always been thought of as aimed at modifying productivity for given conventional factors of production (see, e.g., the tradition started by Griliches (1979)). Our goal is thus to assess the role of R&D in determining the evolution of firm-level productivity over time and the heterogeneity in productivity across firms.

Assume then that the firm spends on R&D and related activities in trying to improve its productivity. Hence, productivity must be considered the product of 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

$$V(k_{jt}, \omega_{jt}) = \max_{i_t, r_t} \pi(k_{jt}, \omega_{jt}) - c_i(i_{jt}) - c_r(r_{jt}) + \beta E[V(k_{jt+1}, \omega_{jt+1})|k_{jt}, \omega_{jt}, i_{jt}, r_{jt}],$$

where $\pi(\cdot)$ denotes per-period profits and $\beta$ is the discount factor. 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 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 main difference between the two types of investments is that they affect the evolution of different state variables, i.e., either the capital stock $k_{jt}$ or the productivity $\omega_{jt}$ of the firm.

The Markovian assumption implies

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

That is, actual productivity $\omega_{jt}$ in period $t$ can be decomposed into expected productivity $g(\omega_{jt-1}, r_{jt-1})$ and a random shock $\xi_{jt}$. Our key assumption is that the impact of R&D on productivity can be expressed through the dependence of the conditional expectation function $g(\cdot)$ on R&D expenditures. In contrast, $\xi_{jt}$ does not depend on R&D expenditures. This productivity innovation 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 (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 $\xi_{jt}$ 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.

If we consider a \textit{ceteris paribus} increase in R&D expenditures that changes \(\omega_{jt}\) to \(\tilde{\omega}_{jt}\), then \(\tilde{\omega}_{jt} - \omega_{jt}\) approximates the effect of this change in productivity on output in percentage terms, i.e., \((Y_{jt} - Y_{jt})/Y_{jt} = \exp(\tilde{\omega}_{jt} - \omega_{jt}) - 1 \approx \tilde{\omega}_{jt} - \omega_{jt}\). That is, the change in \(\omega_{jt}\) shifts the production function and hence measures the change in total factor productivity. Also \(g(\cdot)\) and \(\xi_{jt}\) can be interpreted in percentage terms and decompose the change in total factor productivity. Finally, \(\frac{\partial \omega_{jt}}{\partial \sigma_{jt-1}} = \frac{\partial}{\partial \sigma_{jt-1}} g(\omega_{jt-1}, r_{jt-1})\) is the elasticity of output with respect to R&D expenditures.

A nice feature of our setting is that it encompasses as a particular case the knowledge capital model (see Griliches (1979, 2000)). The impact of R&D on productivity has been traditionally analyzed using this model. A conventional Cobb-Douglas production function is augmented by including the log of knowledge capital \(c_{jt}\) as an extra input, thus yielding

\[
y_{jt} = \beta_0 + \beta_1 k_{jt} + \beta_q h_{jt} + \varepsilon c_{jt} + e_{jt},
\]

where \(\varepsilon\) 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 \(\delta\). Hence, its law of motion can be written as \(C_{jt} = (1 - \delta)C_{jt-1} + R_{jt-1} = C_{jt-1}(1 - \delta + \frac{R_{jt-1}}{C_{jt-1}})\). Taking logs of \(C_{jt}\) we have \(\varepsilon c_{jt} \approx \varepsilon c_{jt-1} + \varepsilon \left(\frac{R_{jt-1}}{C_{jt-1}} - \delta\right)\), where \(\frac{R_{jt-1}}{C_{jt-1}}\) is the rate of investment in knowledge. Letting \(\omega_{jt} = \varepsilon c_{jt}\) it is easy to see that \(\omega_{jt} = \omega_{jt-1} + \varepsilon \left(\frac{\exp(r_{jt-1})}{\exp(\omega_{jt-1}/\varepsilon)} - \delta\right)\) and hence \(\omega_{jt} = g(\omega_{jt-1}, r_{jt-1})\). That is, the “classical” accumulation of knowledge capital induces a particular expression for the conditional expectation function \(g(\cdot)\) that depends on both productivity and R&D expenditures in the previous period.

The knowledge capital model is especially unrealistic in one respect: R&D is likely to produce different outcomes according to chance and hence the accumulation of improvements to productivity is likely to be subjected to shocks. To capture this assume that the effect of the rate of investment in knowledge has an unpredictable component \(\xi_{jt}\). The law of motion becomes \(C_{jt} = C_{jt-1}(1 - \delta + \frac{R_{jt-1}}{C_{jt-1}} + \frac{1}{\varepsilon} \xi_{jt})\). This simple extension causes the law of motion of productivity to be \(\omega_{jt} = g(\omega_{jt-1}, r_{jt-1}) + \xi_{jt}\), which turns out to be our controlled first-order Markov process. Therefore, a useful way to think of our model is as a generalization of the knowledge capital model to the more realistic situation of uncertainty in the R&D process.

At the same time our setting overcomes other problems of the knowledge capital model, in particular the linear accumulation of knowledge from period to period in proportion to R&D expenditures and the linear depreciation. The absence of functional form restrictions on the combined impact of R&D and already attained productivity on future productivity is an important step in the direction of relaxing all these assumptions. Of course, there is a basic difference between the two models. In case of a knowledge capital model, given data on R&D and a guess for the initial condition, it is possible to construct the stock of knowledge capital at all times and apparently fully control for the impact of R&D on productivity. In our setting, in contrast, the random nature of accumulation and the unspecified form of
the law of motion prevents the construction of the “stock of productivity,” which remains unobserved.

3 Estimation strategy

Our model relaxes the assumption of an exogenous Markov process for $\omega_{jt}$. As stressed in Ackerberg, Benkard, Berry & Pakes (2005) making this process endogenous 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. Since all quantities depend on all prices, 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. Second, the absence of data on R&D implies that a critical determinant of the distribution of $\omega_{jt}$ given $\omega_{jt-1}$ is missing and this can create a difficulty for recovering $\omega_{jt}$ from $k_{jt}$, $i_{jt}$, and their lags, the key step in OP.

Buettner (2005) has extended 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 $\omega_{jt}$. The restriction to an index excludes the possibility that current productivity and R&D expenditures affect future productivity in qualitatively different ways. Under certain assumption 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) further notes, however, that there are problems with identification even when data on R&D is available.

The estimation procedure that this paper proposes solves entirely the identification problem when there is data on R&D by using a known function $h(\cdot)$ that is derived from the demand for variable inputs such as labor and materials in order to recover unobserved productivity. Our approach does not rely on an index and frees up the relationship between current productivity, R&D expenditures, and future productivity. It can also solve potentially the identification problem when there is no data on R&D but this point needs further research.\(^4\)

In what follows we first show how the demand for labor can be used to control for unobserved productivity even if the Markov process is shifted by R&D expenditures. We also explain how our proposal to estimate the parameters of the production function nests as a particular case the model proposed by Blundell & Bond (2000). Next we briefly explore the consequences of being unable to perfectly recover unobserved productivity. Then we detail how labor and materials can be used together to control for unobserved productivity, thereby

\(^4\)Muendler (2005) suggests to use investment in physical capital interacted with sector-specific competition variables to proxy for endogenously evolving productivity. His rationale is that firms make R&D decisions in light of their expectations about future market prospects. Hence, in the absence of data on R&D, these competition variables should to some extent capture the drivers of R&D decisions.
increasing the efficiency of the estimator. We finally discuss how imperfect competition can be taken into account and the likelihood of sample selection problems.

### 3.1 Controlling for unobserved productivity

OP suggest using observed investment in order to recover unobserved productivity. Subsequent proposals have focused on input demands. LP use nonparametric methods to invert the demand for materials and substitute for productivity in a gross output production function. Ackerberg, Caves & Frazer (2005) propose using the demands for either labor or materials in a first stage to nonparametrically estimate the random shocks $e_{jt}$ and, in a second stage, use these estimates to back out the Markov process.

Using the demand for labor to control for unobserved productivity is natural in the value added case as is using the demands for either labor or materials in the gross output case. These variable inputs are chosen with current productivity known, and therefore contain information about it. We differ from the previous literature in that we exploit the fact that, given a parametric specification of the production function, the functional form of the input demand functions and their inverses is known. For the sake of simplicity we focus below on the value added case. The extension to the gross output case is straightforward.

Given the Cobb-Douglas production function

$$y_{jt} = \beta_0 + \beta_1 l_{jt} + \beta_2 k_{jt} + \omega_{jt} + e_{jt},$$

the assumption that the firm chooses labor based on the expectation $E(e_{jt}) = 0$ gives the demand for labor as

$$l_{jt} = \frac{1}{1 - \beta_1} \left( \beta_0 + \ln \beta_1 + \beta_2 k_{jt} + \omega_{jt} - (w_{jt} - p_{jt}) \right). \tag{2}$$

Solving for $\omega_{jt}$ we obtain the inverse labor demand function

$$h(l_{jt}, k_{jt}, w_{jt} - p_{jt}) = \lambda_0 + (1 - \beta_1) l_{jt} - \beta_2 k_{jt} + (w_{jt} - p_{jt})$$

where $\lambda_0$ combines the constant terms $-\beta_0$ and $-\ln \beta_1$ and $(w_{jt} - p_{jt})$ is the relative wage (homogeneity of degree zero in prices). 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 $\omega_{jt}$ in the production function cancels out parameters of interest and leaves us with the marginal productivity condition for profit maximization, i.e.,

$$\ln \beta_1 + (y_{jt} - l_{jt}) = w_{jt} - p_{jt} + e_{jt}.$$  

Using its value in period $t - 1$ in our controlled Markov process, however, we have

$$y_{jt} = \beta_0 + \beta_1 l_{jt} + \beta_2 k_{jt} + g(h(l_{jt-1}, k_{jt-1}, w_{jt-1} - p_{jt-1}), r_{jt-1}) + \xi_{jt} + e_{jt}. \tag{3}$$

Only $l_{jt}$ is correlated with $\xi_{jt}$ (since $\xi_{jt}$ is part of $\omega_{jt}$ and $l_{jt}$ is a function of $\omega_{jt}$). Both $k_{jt}$, whose value is determined in period $t - 1$ by $i_{t-1}$, and $r_{jt-1}$ are uncorrelated by virtue of our timing assumptions.

Apart from the presence of R&D expenditures, our estimation equation (3) is seen to
be similar in structure to the second equation of OP and LP when viewed through the lens of Wooldridge’s (2004) GMM framework. The differences between our approach and the existing literature are twofold. First, the parametric specification of the production function implies a known form for the inverse labor demand function $h(\cdot)$ that can be used to control for unobserved productivity. This contrasts with the nonparametric methods in OP and LP. As a consequence, only the conditional expectation function $g(\cdot)$ is unknown and must be estimated nonparametrically. Second, in principle, we need firm-level wage and price data to estimate the model. Below we discuss first identification and then estimation.

Our estimation equation (3) 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 this restriction aids identification, suppose to the contrary that $h(\cdot)$ were 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 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) and also Ackerberg, Caves & Frazer (2005) for an application to the OP/LP framework). Were $h(\cdot)$ of unknown form, the identification condition is violated. To see this, recall that $k_{jt} = (1 - \delta)k_{jt-1} + i(k_{jt-1}, \omega_{jt-1})$ 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 $\omega_{jt-1}$ is by construction a function of all arguments of $h(\cdot)$, thereby making $k_{jt}$ perfectly predictable from the variables in the nonparametric part. Of course, in our setting the inverse labor demand function $h(\cdot)$ is of known form. The central question thus becomes whether $k_{jt}$ is perfectly predictable from the value of $h(\cdot)$ (as opposed to its arguments) and $r_{jt-1}$. Since $h_{jt-1}$ is identical to $\omega_{jt-1}$, we have to ask if $k_{jt-1}$ and hence $k_{jt}$ (via $i(k_{jt-1}, \omega_{jt-1})$) can be inferred from $r_{jt-1}$. 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 $k_{jt-1}$.

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. Lastly, we note that anything that shifts over time the costs of the investments in physical capital and knowledge guarantees identification. The price of equipment goods is likely to vary, for example, and the marginal cost of investment in knowledge depends

---

5In our setting the first equation is the marginal productivity condition for profit maximization. Combining it with our estimating equation (3) may help to estimate the labor coefficient.

6The model remains identified, however, if the log of relative wage is replaced by a set of dummies. This could be an appropriate solution in the absence of wage and price data if the industry can be considered perfectly competitive.
greatly on the nature of the undertaken project. Using \(x_{jt}\) to denote these shifters, the policy functions become

\[ i(k_{jt}, \omega_{jt}, x_{jt}) \text{ and } r(k_{jt}, \omega_{jt}, x_{jt}) \]

Obviously, \(x_{jt}\) cannot be perfectly predicted from \(h_{jt-1}\) and \(r_{jt-1}\). This breaks the functional relationship between \(k_{jt} = (1 - \delta)k_{jt-1} + i(k_{jt-1}, \omega_{jt-1}, x_{jt})\) and \(h_{jt-1}\) and \(r_{jt-1}\).\(^7\)

As suggested by Wooldridge (2004), using a series estimator for the conditional expectation function \(g(\cdot)\) the problem can be cast in the nonlinear GMM framework

\[
E[z_{jt}'(\xi_{jt} + e_{jt})] = E[z_{jt}'v_{jt}(\theta)] = 0,
\]

where \(z_{jt}\) is a vector of instruments and we write the error term \(v_{jt}(\cdot)\) as a function of the parameters \(\theta\) to be estimated. Clearly we need an instrument for \(l_{jt}\). Nonlinear functions of the other variables can be used as instruments, as can be lagged values of \(l_{jt}\) and the other variables. If firms can be assumed to be perfectly competitive, then current wages and prices are exogenous and constitute the most adequate instruments (since demand for labor is directly a function of current wages and prices). We show below how the model specification and the available instruments must be adjusted if there is imperfect competition in the product market and prices are endogenous.

It is worth noting that our model nests, as a particular case, the model proposed by Blundell & Bond (2000). Suppose the Markov process is simply an autoregressive process that does not depend on R&D expenditures so that we have \(g(\omega_{jt-1}) = \rho \omega_{jt-1}\). Using the marginal productivity condition for profit maximization to substitute \(\rho y_{jt-1}\) 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 assumption on the Markov process and the strategy of estimation. In the tradition of OP and LP our method basically proposes the replacement of unobservable autocorrelated productivity by an expression in terms of observed variables and an unpredictable component, whereas their method models the same term through the use of lags of the dependent variable (see Ackerberg, Caves & Frazer (2005) for a detailed description of these two literatures).

### 3.2 Unobservables and misspecification

In practice the inverse labor demand function may allow us to recover \(\omega_{jt}\) at most up to an error term \(u_{jt}\). That is,

\[ h_{jt} + u_{jt} = \omega_{jt}, \]

where the error term \(u_{jt}\) represents all unobservable influences on input demand that we are not dealing with in other ways. Under certain assumptions we are able to work out the consequences of this type of misspecification for the estimates of the labor and capital coefficients \(\beta_l\) and \(\beta_k\) and the conditional expectation function \(g(\cdot)\). Specifically, we

\(^7\)Depending on the construction of the capital stock in the data, we may also be able to account for uncertainty in the impact of investment in physical capital. But once an error term is added to the law of motion for physical capital, \(k_{jt}\) can no longer be written as a function of \(h_{jt-1}\) and \(r_{jt-1}\), and identification is restored.
assume that \( u_{jt} \) is independent of all variables in the nonparametric part of the model
\( g(h(l_{jt}, k_{jt}, w_{jt} - p_{jt}, r_{jt})) \) and their lags and uncorrelated with the random shock \( e_{jt} \).

Without a doubt, these assumptions are restrictive. Their benefit, however, is that they present the simplest case for analyzing the consequences of unobservables and thus set the stage for more realistic cases and our test for misspecification. Before giving details, we sketch a scenario that may be consistent with our assumptions.

**Example 1 (Measurement error)** Suppose that the firm chooses the amount of labor
\( l_{jt}^* \) according to equation (2), but that statistical reporting is subject to measurement error
\( u_{jt} = l_{jt} - l_{jt}^* \). Such errors are common with labor: The amount of labor as reported on a given date is only a proxy for the weighted average of the amount of labor over the relevant production period. In this case it may be reasonable to assume that the measurement error \( u_{jt} \) is independent of the observed amount of labor \( l_{jt} \). Rewriting \( u_{jt} = l_{jt} - l_{jt}^* \) in terms of the inverse labor demand function \( h(\cdot) \) yields \( h_{jt} - (1 - \beta_l)u_{jt} = \omega_{jt} \), where the error term \(-(1 - \beta_l)u_{jt}\) satisfies our assumptions.

We are still able to consistently estimate the labor and capital coefficients \( \beta_l \) and \( \beta_k \) but not the conditional expectation function \( g(h_{jt-1} + u_{jt-1}, r_{jt-1}) = E[\omega_{jt}|h_{jt-1} + u_{jt-1}, r_{jt-1}] \). Instead, we estimate the expectation (with respect to the error term) of the conditional expectation function as given by

\[
\tilde{g}(h_{jt-1}, r_{jt-1}) = E[\omega_{jt}|h_{jt-1}, r_{jt-1}]
\]

\[
= E[E[\omega_{jt}|h_{jt-1} + u_{jt-1}, r_{jt-1}]] = E[E[\omega_{jt}|h_{jt-1} + u_{jt-1}, r_{jt-1}]|h_{jt-1}, r_{jt-1}].
\]

(4)

In general \( g(\cdot) \neq \tilde{g}(\cdot) \), but the two functions become more similar as the error term becomes less important. The function \( \tilde{g}(\cdot) \) averages out over the heterogeneity across firms as expressed in \( u_{jt} \) and, depending on the application at hand, may be as much of interest as the function \( g(\cdot) \). Moreover, given that the error form of the Markov process is \( \omega_{jt} = g(h_{jt-1} + u_{jt-1}, r_{jt-1}) + \xi_{jt} \), we now have \( \omega_{jt} = \tilde{g}(h_{jt-1}, r_{jt-1}) + \xi_{jt} \), where by the law of iterated expectations the productivity innovation \( \xi_{jt} \) is mean independent of all variables in \( \tilde{g}(l_{jt-1}, k_{jt-1}, w_{jt-1} - p_{jt-1}, r_{jt-1}) \) and their lags. Hence, we can continue to instrument as before. A formal proof is given in Appendix A.\(^8\)

We can further extend our treatment of unobservables to allow for linear correlation between \( u_{jt} \) and some or all of the variables in \( h(l_{jt}, k_{jt}, w_{jt} - p_{jt}) \). To simplify the exposition, in what follows we maintain the assumption that the conditional expectation \( E(u_{jt}|l_{jt}, k_{jt}, w_{jt} - p_{jt}) = \gamma_0 + \gamma_1 l_{jt} \) is linear in \( l_{jt} \) with \( \gamma_1 > 0 \). Then we have \( u_{jt} = \gamma_0 + \gamma_1 l_{jt} + v_{jt} \), where we assume as before that the error term \( v_{jt} \) is independent of all variables in the nonparametric part of the model and their lags and uncorrelated with the

\(^8\)In the context of a gross output production function we are likely to have one error term per input. If these error terms can be considered independent of all relevant variables (in particular of both lagged inputs), then lags of the input other than the one whose inverse demand function is being used are valid instruments. The error terms may be correlated with each other.
random shock $\epsilon_{jt}$. The price to be paid for accommodating this more general relationship between observables and unobservables is that we are no longer able to recover the structural parameters in the inverse labor demand function and possibly also in the production function. We sketch two scenarios.

**Example 2 (Optimization error and omitted variable)** Suppose that the firm perceives its productivity to be $\omega_{jt} + u_{jt}$ instead of $\omega_{jt}$. Replacing $\omega_{jt}$ by $\omega_{jt} + u_{jt}$ in equation (2) and solving for $\omega_{jt}$ yields $\omega_{jt} = h_{jt} - u_{jt}$, where $l_{jt}$ and $u_{jt}$ are linearly correlated (because, by virtue of equation (2), $l_{jt}$ is chosen as a linear function of $u_{jt}$). Hence, using $u_{jt} = \gamma_0 + \gamma_1 l_{jt} + v_{jt}$, we have

$$\omega_{jt} = h_{jt} - (\gamma_0 + \gamma_1 l_{jt}) - v_{jt},$$

where the error term $-v_{jt}$ satisfies our assumptions. Combining terms, however, shows that the coefficient of $l_{jt}$ on the RHS of the above equation has changed from $1 - \beta_l$ to $1 - \beta_l - \gamma_1$. Hence, the labor coefficient is biased upward in the inverse input demand function.

A similar situation may arise in the absence of wage and price data: If the industry cannot be considered perfectly competitive, then the firm-specific relative wage in equation (2) must be replaced by the sum of a dummy variable and an error term that captures unobserved idiosyncratic deviations from the average relative wage.

**Example 3 (Measurement error)** The classical “error in variable” assumption is $l_{jt} = l_{jt}^* + u_{jt}$ with $u_{jt}$ being uncorrelated with $l_{jt}^*$ and hence correlated with $l_{jt}$. Recall that before we looked at the opposite extreme. Rewriting $l_{jt} = l_{jt}^* + u_{jt}$ in terms of the inverse labor demand function $h(\cdot)$, using $u_{jt} = \gamma_0 + \gamma_1 l_{jt} + v_{jt}$, and combining terms shows that the coefficient of $l_{jt}$ has changed from $(1 - \beta_l)$ to $(1 - \beta_l)(1 - \gamma_1)$. Note that at the same time the coefficient of $l_{jt}$ in the production function is changing from $\beta_l$ to $\beta_l - \gamma_1$. Hence, the labor coefficient is biased upward in the inverse input demand function but, in contrast to Example 2, biased downward in the production function.

Provided we refrain from imposing equality of the structural parameters in the production function and the inverse labor demand function, we are still able to consistently estimate $\beta_l$ and $\beta_k$ (from the production function alone) as well as $\tilde{g}(\cdot)$ in Example 2 but not Example 3. Both examples stress, however, that a correlated unobservable causes the parameters in the inverse labor demand function to diverge from their counterparts in the production function. This forms the basis for our test for misspecification. By testing the null hypothesis that the structural parameters in the two parts of the model are equal, we may rule out that a correlated unobservable is severely diminishing our ability to back out unobserved productivity and that our model is therefore misspecified. Of course, our test cannot detect the presence of an independent unobservable, but this case is much less harmful for our estimator in the first place. Note that our test is especially useful in picking up errors in measuring labor (or materials the gross output case) because they bias the
parameters in the two parts of the model in opposite directions as illustrated by Example 3. It is probably not as powerful when it comes to errors in measuring capital because they will tend to attenuate the capital coefficient in both the parts of the model.

### 3.3 Combining multiple input demands

If a gross output production function is used, both the demand for labor and the demand for materials can be inverted to back out unobserved productivity. In this case we have two inverse input demand functions which, ignoring for simplicity the specific form of the constant, are given by

\[
h_l(l_{jt}, k_{jt}, w_{jt} - p_{jt}, p_{Mjt} - p_{jt}) = \lambda_{0l} + (1 - \beta_l - \beta_m)l_{jt} - \beta_k k \\
+ (1 - \beta_m)(w_{jt} - p_{jt}) + \beta_m(p_{Mjt} - p_{jt}),
\]

\[
h_m(m_{jt}, k_{jt}, w_{jt} - p_{jt}, p_{Mjt} - p_{jt}) = \lambda_{0m} + (1 - \beta_l - \beta_m)m_{jt} - \beta_k k \\
+ \beta_l(w_{jt} - p_{jt}) + (1 - \beta_l)(p_{Mjt} - p_{jt}),
\]

where \((p_{Mjt} - p_{jt})\) is the relative price of materials. While the first expression is a function of capital and labor and the second of capital and materials, the relative prices for both inputs enter both expressions although with different coefficients.

This suggests that the two inverse input demand functions should be used to estimate the parameters of the production function and preferably combined to increase efficiency. To do this the specification must be made more flexible. If both functions are identical to \(\omega_{jt}\), then they have to be equal to each other. But efficiency could not increase if they were bringing exactly the same information. Moreover, even if this could be acceptable in theory, the data are going to refute it.

The problem is easily solved by assuming that, as is likely to be the case in practice, the inverse input demand functions are identical to each other and to \(\omega_{jt}\) up to the error terms \(u_{ljt}\) and \(u_{mjt}\). That is,

\[
h_{ljt} + u_{ljt} = h_{mjt} + u_{mjt} = \omega_{jt},
\]

where the errors terms \(u_{ljt}\) and \(u_{mjt}\) represent all unobservable influences on input demand that we are not dealing with in other ways. We assume that \(u_{ljt}\) and \(u_{mjt}\) are independent of all variables in the nonparametric part of the model and their lags and uncorrelated with the random shock \(e_{jt}\). Then, by the law of iterated expectations, we have

\[
E[g(h_{ljt-1} + u_{ljt-1}, r_{jt-1})|h_{ljt-1}, r_{jt-1}] = \tilde{g}_l(h_{ljt-1}, r_{jt-1}),
\]

\[
E[g(h_{mjt-1} + u_{mjt-1}, r_{jt-1})|h_{mjt-1}, r_{jt-1}] = \tilde{g}_m(h_{mjt-1}, r_{jt-1}).
\]
and the system of estimation equations is

\[
\begin{align*}
y_{jt} &= \beta_0 + \beta l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + \tilde{g}_l(h_{ljt-1}, r_{jt-1}) + \tilde{\xi}_{ljt} + e_{jt}, \\
y_{jt} &= \beta_0 + \beta l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + \tilde{g}_m(h_{mjt-1}, r_{jt-1}) + \tilde{\xi}_{mjt} + e_{jt}.
\end{align*}
\]

This problem can again be cast in the nonlinear GMM framework. Combining multiple input demands yields two types of efficiency gains. First, we can increase efficiency by restricting the structural production function parameters to be the same in both equations. In practice, it is prudent to first ensure that this restriction is supported by the data. Having estimated the equations separately, it is convenient to use a Wald test of the null hypothesis of equality of (one or more) parameters. Second, we can increase efficiency by exploiting the correlation in the error terms of both equations. Of course, to realize this last gain we have to use the optimal GMM weighting matrix. In the extreme case of both equations containing exactly the same information, this matrix has less than full rank. More generally, if this matrix is “close” to singular, then the gain is small and it may not be worthwhile to combine multiple input demands.

### 3.4 Imperfect competition

Until now we have 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). In what follows we thus assume that firms set the price of their output according to its cost and the specific demand conditions they are facing. Demand conditions include the particular type of equilibrium which prevails in the market (which we do not observe and assume to be constant over time) and the demand shifters, i.e., the values of variables that displace demand given price.

Returning for the sake of simplicity to the value added case, we specifically assume that firms face 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 evaluated at the equilibrium price and the particular value of the demand shifter and, for convenience, is written as a function of \( p_{jt} = \ln P_{jt} \) and \( z_{jt} = \ln Z_{jt} \). 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 inverse labor demand function

\[
h_{IC}(l_{jt}, k_{jt}, w_{jt} - p_{jt}, p_{jt}, z_{jt}) = \lambda_0 + (1 - \beta l_{jt} - \beta_k k_{jt} + (w_{jt} - p_{jt}) - \ln \left(1 - \frac{1}{\eta(p_{jt}, z_{jt})}\right).
\]
Thus, the estimation equations is

\[ y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + g \left( h_{jt-1} - \ln \left( 1 - \frac{1}{\eta(p_{jt-1}, z_{jt-1})} \right), r_{jt-1} \right) + \xi_{jt} + e_{jt}. \]

Using series estimators for the unknown functions \( g(\cdot) \) and \( \eta(\cdot) \), this problem can again be cast in the nonlinear GMM framework. As both \( p_{jt} \) and \( z_{jt} \) enter the equations lagged they are expected to be uncorrelated with the productivity innovation \( \xi_{jt} \). Moreover, testing for imperfect competition can be done by adding an unknown function in \( p_{jt-1} \) and \( z_{jt-1} \) to \( h_{jt-1} \) inside the conditional expectation function.

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

### 3.5 Sample selection

A potential problem in the estimation of production functions is sample selection. If a firm’s dynamic programming problem generates an optimal exit decision, based on the comparison between the sell-off value of the firm and its expected profitability in the future, then this decision is a function of current productivity. The simplest model, based on an exogenous Markov process, predicts that if an adversely enough shock to productivity is followed immediately by exit, then there will be a negative correlation between the shocks and the capital stocks of the firms that remain in the industry. Hence, sample selection will lead to biased estimates.

Accounting for R&D expenditures in the Markov process complicates matters. On the one hand, a firm now has an instrument to try to rectify an adverse shock and the optimal exit decision is likely to become more complicated. To begin with, there are many more relevant decisions such as beginning, continuing, or stopping innovative activities whilst remaining in the industry, and exiting in any of the different positions. On the other hand, a firm now is more likely to remain in the industry despite an adverse shock. Innovative activities often imply large sunk cost which will make the firm more reluctant to exit the industry or at least to exit it immediately. This will tend to alleviate the selection problem. At this stage we do not model any of these decisions. Instead, we simply explore whether there is a link between exit decisions and estimated productivity.

### 4 Econometric specification

Below we explain the details of the model specification, the estimation procedure and the tests that we apply to the different specifications.

Our preliminary estimates indicate that in some industries it is useful to add a time
trend so that the production function becomes

\[ y_{jt} = \beta_0 + \beta_t t + \beta_l l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + \omega_{jt} + e_{jt}. \]

In these cases the inverse input demand function \( h(\cdot) \) contains, in addition to the constant, a term equal to \(-\beta_t t\). One can say that there is an “observable” trend in the evolution of productivity that is treated separately from \( \omega_{jt} \) but of course taken into account when substituting \( h_{jt} \) for \( \omega_{jt} \). Our goal is thus to estimate

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

where the conditional expectation function \( g(\cdot) \) is of unknown form and thus must be estimated nonparametrically.

As suggested by Wooldridge (2004) when modeling an unknown function \( q(v, u) \) of two variables \( v \) and \( u \) we use a series estimator made of a “complete set” of polynomials of degree \( Q \) (see Judd 1998), i.e., all polynomials of the form \( v^j u^k \), where \( j \) and \( k \) are nonnegative integers such that \( j + k \leq Q \). When the unknown function \( q(\cdot) \) has a single argument, we use a polynomial of degree \( Q \) to model it, i.e., \( q(v) = \rho_0 + \rho_1 v + \ldots + \rho_Q v^Q \).

Taking into account that there are firms that do not perform R&D, the most general formulation is

\[ y_{jt} = \beta_0 + \beta_t t + \beta_l l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + 1(R_{jt-1} = 0) g_0(h_{jt-1}) + 1(R_{jt-1} > 0) g_1(h_{jt-1}, r_{jt-1}) + \xi_{jt} + e_{jt}. \] (5)

This allows for a different unknown function when the firm adopts the corner solution of zero R&D expenditures and when it chooses positive R&D expenditures.

It is important to note that any constant that its arguments may have will be subsumed in the constant of the unknown function. Our specification is therefore

\[
\begin{align*}
g_0(h_{jt-1}) &= g_{00} + g_{01}(h_{jt-1} - \lambda_0), \\
g_1(h_{jt-1}, r_{jt-1}) &= g_{10} + g_{11}(h_{jt-1} - \lambda_0, r_{jt-1}),
\end{align*}
\]

where in \( g_{00} \) and \( g_{10} \) we collapse the constants of the unknown functions \( g_0(\cdot) \) and \( g_1(\cdot) \) and the constant of \( h_{jt-1} \). Note that \( g_{00} \), \( g_{10} \), and \( \beta_0 \) cannot be estimated separately. We thus estimate the constant for nonperformers \( g_{00} \) together with the constant of the production function \( \beta_0 \) and include a dummy for performers to measure the difference between constants \( \beta_0 + g_{10} - (\beta_0 + g_{00}) = g_{10} - g_{00} \).

When we combine the demands for labor and materials, we allow \( \tilde{g}_l(\cdot) \) and \( \tilde{g}_m(\cdot) \) to be completely different unknown functions (including constants) and model them as described above. Note that while \( \tilde{g}_l(\cdot) \) and \( \tilde{g}_m(\cdot) \) are expectations of the same underlying conditional expectation function \( g(\cdot) \), these expectations are computed from the potentially different
distributions of the error terms $u_{jt}$ and $u_{mjt}$.

In the case of imperfect competition, where we have to nonparametrically estimate the absolute value of the elasticity of demand, we impose the theoretical restriction that $\eta(\cdot) > 1$ by using the specification $\eta(p_{jt-1}, z_{jt-1}) = 1 + \exp(q(p_{jt-1}, z_{jt-1}))$, where $q(\cdot)$ is modeled as described above.

As advanced above, we interact a vector of instruments with the errors terms of the estimation equations, specified in terms of series estimators, and base the estimation of the parameters $\theta$ on the nonlinear GMM problem

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

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. The generalization to the case in which we specify more than one equation per firm is straightforward (see, e.g., Wooldridge (1996) for systems of equations). 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 $\theta$ and then we compute the optimal estimator which uses weighting matrix $A_N = \left( \frac{1}{N} \sum_j z_j' v_j(\hat{\theta}) v_j(\hat{\theta})' z_j \right)^{-1}$.

Let us detail the instruments that we are interacting with the error terms. As discussed before, $k_{jt}$ is always a valid instrument because it is not correlated with $\xi_{jt}$ because the latter is unpredictable when $i_{jt-1}$ is chosen. Labor and materials, however, are contemporaneously correlated with the innovation to productivity. The lags of these variables are valid instruments but when the demand for one of these inputs is being used to substitute for $\omega_{jt}$ it appears itself in $h_{jt-1}$. We can use the lag of the other. Constant and trend are valid instruments. To estimate the constant, the trend, and capital, labor, and materials coefficients we have therefore for the moment four instruments. This leaves us with the need for at least one more instrument. We use as instruments for the whole equation the complete set of polynomials of degree $Q$ in the variables which enter $h_{jt-1}$, the powers up to degree $Q$ of $r_{jt-1}$, and the interactions up to degree $Q$ of the variables which enter $h_{jt-1}$ and $r_{jt-1}$. The nonlinear functions of all exogenous variables included in these polynomials provide enough instruments.

In practice we set $Q = 3$ and use polynomials of order three. Hence, when there are four variables in the inverse input demand function $h(\cdot)$, say $l_{jt-1}$, $k_{jt-1}$, $w_{jt-1} - p_{jt-1}$, and $pM_{jt-1} - p_{jt-1}$, we use as instruments the polynomials which result from the complete set of polynomials of degree 3 corresponding to the third power of $h_{jt-1}$ (34 instruments), plus 3 terms which correspond to the powers of $r_{jt-1}$ (3 instruments) and 12 interactions formed from the products $h_{jt-1}r_{jt-1}$, $h_{jt-1}^2 r_{jt-1}$, and $h_{jt-1}r_{jt-1}^2$ (12 instruments). In fact when we enter $p_{jt-1}$ linearly we use it detached from the other prices and we also need a dummy for the firms that perform R&D (2 more instruments). In addition, when there are enough degrees of freedom we instrument separately $h_{jt-1}$ for nonperformers and $h_{jt-1}$ and
r_{jt-1} for performers by interacting the instruments with the dummy for performers. And we have the exogenous variables included in the equation: constant, trend, current capital and lagged materials (4 instruments). This gives a total of 34 + 34 + 12 + 1 + 2 + 4 = 90 instruments. When we combine the demands for labor and materials, both equations have the same number of instruments (recall that not all are equal) and hence we have a total of 190 instruments.

It is worth noting that, given these instruments, the estimator that we apply has exactly the form of the GMM version of Ai & Chen’s (2003) sieve minimum distance estimator, a nonparametric least squares technique (see Newey & Powell 2003). This means that, if the conditional expectation function \( g(\cdot) \) is specified in terms of variables which are correlated with the error term of the estimation equation, we still obtain a consistent and asymptotically normal estimator of the parameters by specifying the instrumenting polynomials in terms of exogenous conditioning variables.

The value of the GMM objective function for the optimal estimator, multiplied by \( N \), has a limiting \( \chi^2 \) distribution with \( L - P \) degrees of freedom, where \( L \) is the number of instruments and \( P \) the number of parameters to be estimated.\(^9\) We use it as a test of overidentifying restrictions or validity of the moment conditions based on the instruments.

We test whether the model satisfies certain restrictions by computing the restricted estimator using the weighting matrix for the optimal estimator and then comparing the values of the properly scaled objective functions. The difference has a limiting \( \chi^2 \) distribution with degrees of freedom equal to the number of restrictions. The tests for misspecification and imperfect competition are of this form. In addition, we test whether R&D plays a role in the conditional expectation function, whether this function is separable, and whether it is consistent with the knowledge capital model. In the first case, we test whether all terms in \( r_{jt-1} \) can be excluded from the conditional expectation function \( g_{11}(h_{jt-1} - \lambda_0, r_{jt-1}) \) for performers plus the equality of the common part of the conditional expectation functions for performers and nonperformers, i.e., \( g_{11}(h_{jt-1} - \lambda_0, r_{jt-1}) = g_{01}(h_{jt-1} - \lambda_0) \) for all \( r_{jt-1} \). In the second case, we test whether \( g_{11}(h_{jt-1} - \lambda_0, r_{jt-1}) \) can be broken up into two additively separable functions \( g_{11}(h_{jt-1} - \lambda_0) \) and \( g_{12}(r_{jt-1}) \). Finally, we test whether the conditional expectation function satisfies the functional form restrictions implied by the knowledge capital model, i.e., whether \( g_{01}(h_{jt-1} - \lambda_0) = h_{jt-1} - \lambda_0 \) and

\[
g_{11}(h_{jt-1} - \lambda_0, r_{jt-1}) = h_{jt-1} - \lambda_0 + \varepsilon_0 \frac{\exp(r_{jt-1})}{\exp(h_{jt-1} - \lambda_0/\varepsilon_0)}.
\]

(Note that the term \(-\varepsilon \delta\) is absorbed by the constants \( g_{00} \) and \( g_{10} \).)

Once the model is estimated we can compute \( \omega_{jt}, h_{jt}, \) and \( g(\cdot) \) up to a constant. We can also obtain an estimate of \( \xi_{jt} \) up to a constant as the difference between the estimates of \( \omega_{jt} \) and \( g(\cdot) \). Recall that the productivity of firm \( j \) in period \( t \) is given by \( \beta h_t + \omega_{jt} = \beta h_t + g(\omega_{jt-1}, r_{jt-1}) + \xi_{jt} \) with \( \omega_{jt} = h_{jt} \). Using the notational convention that \( \hat{\omega}_{jt}, \hat{h}_{jt}, \) and

\(^9\)Our baseline specification has 18 parameters: constant, trend, three production function coefficients, and thirteen coefficients in the series approximations.
\(g(\cdot)\) represent the estimates up to a constant, we have

\[
\hat{\omega}_{jt} = \hat{h}_{jt} = -\hat{\beta}_t + (1 - \hat{\beta}_t - \hat{\beta}_m)j_{jt} - \hat{\beta}_k k_{jt} + (1 - \hat{\beta}_m)(w_{jt} - p_{jt}) + \hat{\beta}_m(p_{Mjt} - p_{jt})
\]

and

\[
\hat{g}(\hat{h}_{jt-1}, r_{jt-1}) = 1\{R_{jt-1} = 0\}\hat{g}_{01}(\hat{h}_{jt-1}) + 1\{R_{jt-1} > 0\}\{g_{10} - g_{00}\} + \hat{g}_{11}(\hat{h}_{jt-1}, r_{t-1})].
\]

This implies that we can estimate \(Var(\omega_{jt}), Var(g(\cdot))\) and \(Var(\xi_{jt})\) as well as \(Cov(g(\cdot), \xi_{jt})\) and the correlation coefficient \(Corr(g(\cdot), \xi_{jt}) = Cov(g(\cdot), \xi_{jt})/\sqrt{Var(g(\cdot))Var(\xi_{jt})}\). We can also estimate the random shocks \(e_{jt}\) and their variance \(Var(e_{jt})\). When we combine multiple input demands, we compute the variances and covariances of \(\omega_{jt}, g(\cdot), \) and \(\xi_{jt}\) from an average of the input-specific estimates.

## 5 Data

We use an unbalanced panel of Spanish manufacturing firms in nine industries during the 1990s. Our data come from the ESEE (Encuesta Sobre Estrategias Empresariales) survey, a firm-level panel survey of Spanish manufacturing sponsored by the Ministry of Industry.\(^{10}\)

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 the respondents represent a more or less self-selected 70% of all firms of this size. Some firms vanish from the sample, due to both exit and attrition. The two reasons can be distinguished, and attrition remained within acceptable limits. From hereon we reserve the word exit to characterize shutdown by death or abandonment of activity. To preserve representativeness, samples of newly created firms were added to the initial sample every year.

We account for the survey design as follows. First, to compare the productivities of firms that perform R&D to those of firms that do not perform R&D we conduct separate tests on the subsamples of small and large firms. Second, to be able to interpret some of our descriptive statistics as aggregates that are representative for an industry as a whole, we replicate the subsample of small firms \(70 = 14\) times before merging it with the subsample of large firms. Details on industry and variable definitions can be found in Appendix B.

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. Table 1 shows the number of observations and firms by industry. The samples are of moderate size. Firms tend to remain in the sample for short periods, ranging from a minimum of two

\(^{10}\)This data has been used elsewhere, e.g., in Gonzalez, Jaumandreu & Pazo (2005) to study the effect of subsidies to R&D and in Delgado, Farinas & Ruano (2002) to study the productivity of exporting firms.
years to a maximum of 10 years between 1990 and 1999. The descriptive statistics in Table 1 are computed for the period from 1991 to 1999 and exclude the first observation for each firm. The small size of the samples is compensated for by the quality of the data, which seems to keep noise coming from errors in variables at relatively low levels.

Entry and exit reported in Table 1 refer to the incorporation of newly created firms and to 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 exiters (from 5% to above 10% in a few cases).

Table 1 shows that the 1990s were a period of rapid output growth, coupled with stagnant or at best slightly increasing employment and intense investment in physical capital. The growth of prices, averaged from the growth of prices as reported individually by each firm, is moderate.

6 Results

6.1 Innovative activities

The R&D intensity of Spanish manufacturing firms is low by European standards, but R&D became increasingly important during the 1990s (see, for example, European Commission 2001). Table 1 reveals that the nine industries are rather different when it comes to innovative activities of firms. This can be seen along three dimensions, namely the share of firms that perform R&D, the degree of persistence in performing R&D over time, and R&D intensity among performers (defined as the ratio of R&D expenditures to output).

Three industries are highly active: Chemical products (3), agricultural and industrial machinery (4), and transport equipment (6). The share of firms that perform R&D during at least one year in the sample period is two thirds, with slightly more than 40% of stable performers that engage in R&D in all years and slightly more than 20% of occasional performers that engage in R&D in some (but not all) years. Dividing the share of stable performers by the combined share of stable and occasional performers yields the conditional share of stable performers and gives an indication of the persistence in performing R&D over time. With about 65% the degree of persistence is is very high. Finally, the average R&D intensity among performers ranges from 1.5% to 2.1%.

Four industries are in an intermediate position: Metals and metal products (1), non-metallic minerals (2), food, drink and tobacco (7), and textile, leather and shoes (8). The share of performers is lower than one half, but it is near one half in the first two industries. With a conditional share of stable performers of about 40% the degree of persistence tends to be lower. The average R&D intensity among performers is between 0.5% and 0.7% with a much lower value of 0.3% in industry 7.

Two industries, namely timber and furniture (9) and paper and printing products (10), exhibit low innovative activity. The first industry is weak in the share of performers (below 20%), degree of persistence, and R&D intensity. In the second industry the degree of
persistence is somewhat higher with a conditional share of stable performers of 46% but the share of performers remains below 30% and the average R&D intensity is a mere 0.3%.

This heterogeneity in the three dimensions of innovative activities makes it difficult to fit a single model to explain the impact of R&D on productivity. In addition, it is likely that heterogeneity across firms within industries is important because the level of aggregation used in defining these industries encompasses many different specific innovative activities.

### 6.2 Production function and Markov process

Table 2 summarizes different production function estimates. The first three columns report the coefficients estimated from simple OLS regressions of the log of output on the logs of inputs. The coefficients are reasonable as usual when running OLS on logs (but not when running OLS on first-differences of logs), and returns to scale are remarkably close to constancy. The share of capital in value added, as given by the capital coefficient scaled by the sum of the labor and capital coefficients, turns out to be between the pretty usual values of 0.15 and 0.35.

The next six columns of Table 2 report the coefficients estimated when we use the demand for labor to back out unobserved productivity. Specifying the law of motion of productivity to be an exogenous Markov process that does not depend on R&D expenditures yields the coefficients reported in columns four to six. Compared to the simple OLS regressions, the changes go in the direction that is expected from theory. The labor coefficients decrease considerably in all industries while the capital coefficients increase somewhat in 7 industries. The materials coefficients show no particular pattern. Changes are as expected not huge because we are comparing estimates in logs (as opposed to first-differences of logs) for both the exogenous and the controlled Markov process. All this matches the results in OP and LP.

Columns seven to nine show the coefficients obtained when specifying a controlled Markov process. Again, compared to the simple OLS regressions, the changes go in the expected direction. The labor coefficients decrease in 8 cases, the capital coefficients increase in 5 cases and are virtually the same in 2 more cases. In fact, changes from the exogenous to the controlled Markov process seem not to exhibit a distinct pattern. 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 6.3 for details).

*** ADD COMPARISON TO ALTERNATIVE SPECIFICATIONS USING THE DEMAND FOR MATERIALS OR THE DEMANDS FOR BOTH LABOR AND MATERIALS TO BACK OUT UNOBSERVED PRODUCTIVITY (PUT TABLES IN APPENDIX). FIRST, PRELIMINARY ESTIMATES INDICATE THAT THE SIGN OF CHANGES RELATIVE TO OLS DEPENDS ON THE INPUT USED. HOWEVER, OUR CONCLUSIONS REGARDING PRODUCTIVITY LEVELS AND GROWTH REMAIN REMARK-
ABLY STABLE. SECOND, THE CASE FOR COMBINING MULTIPLE INPUT DEMANDS IS NOT STRONG. THE WALD TESTS REJECT THE NULL HYPOTHESIS FOR EQUALITY OF PARAMETERS ACROSS EQUATIONS IN MANY CASES, ESPECIALLY WHEN WE TEST ALL THE STRUCTURAL PRODUCTION FUNCTION PARAMETERS AT THE SAME TIME. (OF COURSE, THE WALD TEST TENDS TO REJECT TOO OFTEN IN SMALL SAMPLES.) MOREOVER, THE OPTIMAL GMM WEIGHTING MATRIX HAS LESS THAN FULL RANK IN INDUSTRIES 2, 4, 6, 9 AND IS “CLOSE” TO SINGULAR IN THE REMAINING INDUSTRIES, INDICATING THAT THE GAIN IN EFFICIENCY IS SMALL. ***

*** ADD COMPARISON TO OP METHOD (PUT TABLES IN APPENDIX). ***

To check the validity of the estimates in Table 2 we have subjected them to a battery of tests. The results are reported in Table 3. We first test for the overidentifying restrictions or validity of the moment conditions based on the instruments as described in Section 4. The test statistic is too high for the usual significance levels in only the case of industry 1. The other values indicate the validity of the moment conditions by a wide margin.

We next turn to our test for misspecification. Recall from Section 3.2 that the presence of a correlated unobservable causes the parameters in the inverse labor demand function to diverge from their counterparts in the production function. Fortunately, while we must reject the null hypothesis of equality in industries 1, 7, and 10, in the remaining industries the test suggests by a wide margin that we may rule out that a correlated unobservable is severely diminishing our ability to back out unobserved productivity and that our model is therefore misspecified.

*** ADD TEST FOR IMPERFECT COMPETITION. ALL THE ESTIMATES REPORTED IN THE PAPER ARE DONE ASSUMING PERFECT COMPETITION. OUR WORK TO DATE SUGGESTS THAT THIS ASSUMPTION IS OFTEN TIMES REJECTED BY THE DATA. PRELIMINARY ESTIMATES ASSUMING IMPERFECT COMPETITION YIELD REASONABLE PRICE ELASTICITIES AND SHOW THAT BASIC RESULTS DO NOT CHANGE WITH IMPERFECT COMPETITION. ***

We next turn to the conditional expectation function $g(\cdot)$ that describes the Markov process of unobserved productivity. We test for the role of R&D by comparing the controlled with the exogenous Markov process. The result is overwhelming: In all cases the constraints imposed by the model with the exogenous Markov process are clearly rejected. We also test whether the conditional expectation function $g_{11}(\cdot)$ for firms that perform R&D is separable in its arguments $\omega_{jt-1}$ and $r_{jt-1}$ (see again Section 4). The test statistics indicate that this is only the case in industries 1, 4, and perhaps 9. From hereon we impose separability on industry 4, where it slightly improves the estimates, but we keep nonseparability in industry 1, where separability does not seem to change anything. Given the limited number of firms that perform R&D in industries 9 and 10, we also impose separability in the interest of parsimony. The main result, however, is 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.

*** ADD TEST FOR KNOWLEDGE CAPITAL MODEL. ***

Since the tests fully validate the model, we use it to compute productivity levels and productivity growth. Below we also use the dependence of the conditional expectation function \( g(\cdot) \) on R&D expenditures to describe in more detail the nature of the R&D process. Since this function plays a central role, however, we first examine its contribution to explaining the data in relation to the other sources of variation.

As explained in Section 4, we can retrieve \( \omega_{jt} \), \( g(\cdot) \), and \( \xi_{jt} \) up to a constant from the production function estimates. The first two of the last four columns of Table 3 tell us the variance of unobserved productivity \( \omega_{jt} \) and the variance of the random shock \( e_{jt} \). 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. The next column gives the ratio of the variance of the productivity innovation \( \xi_{jt} \) to the variance of actual productivity \( \omega_{jt} \). The ratio shows that the unpredictable component accounts for a large part of attained productivity, between 20% and 50%. 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. The last column gives the correlation between expected productivity and the innovation to productivity. Since the correlation is low, this further validates the model. Only in industries 1 and 8 the correlation tends to be a bit higher, as could have already been guessed from the specification test.

### 6.3 Productivity levels

Using our estimate of expected productivity \( \hat{g}(\hat{h}_{jt-1}, r_{jt-1}) \) we compare productivity levels across firms that perform R&D and firms that do not perform R&D. We first compare descriptively the distribution of expected productivity across firms and years without R&D to the distribution of expected productivity across firms and years with R&D. Then we compare formally the means and variances of the distributions and the distributions themselves.

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 (solid line) and without R&D (dashed line). To be able to interpret these descriptive measures as representative aggregates, we proceed as described in Section 5. For each industry Figure 1 shows the density and distribution functions for performers and nonperformers. This exercise yields some interesting insights. In all industries but 4, 9, and 10, the distribution for performers is to the right of the distribution for nonperformers. This strongly suggests stochastic dominance. In contrast, in industries 4 and 10 the distribution functions openly cross: Attaining the highest levels seems more likely for the nonperformers than
for the performers. In industry 9 the distribution for nonperformers dominates the one for 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 heterogeneous innovative activities and hence blurs remarkable differences in the impact of the investment in knowledge on the productivity of firms.

Turning to the moments of the distributions, the difference in means is computed as

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

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

\[
t = \frac{\hat{g}_0 - \hat{g}_1}{\sqrt{Var(g_{01})/(NT_0 - 1) + Var(g_{11})/(NT_1 - 1)}}
\]

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.

Column four of Table 4 reports the difference in means \( \hat{g}_1 - \hat{g}_0 \) (with the opposite sign of the test statistic for the sake of intuition) and the next four columns report the standard deviations and the test statistics along with their probability values. All this is done separately for the subsamples of small and large firms for the reasons given in Section 5. 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 4% and 5% and up to 9%. They are often larger for the smaller firms. In the two industries that exhibit low innovative activity, however, one size group shows a lower mean of expected productivity than the other: The small firms in industry 9 and the large firms in industry 10. The formal statistical test duly rejects, at the usual significance levels, the hypothesis of a higher mean of expected productivity among performers than among nonperformers in these two cases and in both size groups 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 4, it is rejected for both size groups in industries 4, 6, 7, and
10, for small firms in industries 2, 3 and 9, and for large firms in industries 1 and 8.

The above results suggest to compare the distributions themselves. The Kolgomorov-
Smirnov test allows for a comparison of the empirical distributions of independent observations in two independent samples (see Barret & Donald (2003) and Delgado et al. (2002) for similar applications). To apply this test we have to define our samples in a different way. 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 \(F_{N_0}(g) = G_{N_1}(g)\) for all \(g\), i.e., the distributions of expected productivity are equal, and the one-sided test of the hypothesis \(F_{N_0}(g) \leq G_{N_1}(g)\) for all \(g\), 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

\[
S^1 = \sqrt{\frac{N_0 N_1}{N_0 + N_1}} \max_g \{|F_{N_0}(g) - G_{N_1}(g)|\}, \quad S^2 = \sqrt{\frac{N_0 N_1}{N_0 + N_1}} \max_g \{F_{N_0}(g) - G_{N_1}(g)\},
\]

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

When the number of firms is small the tests tend to be inconclusive, so we limit their application to cases in which we have at least 20 performers and 20 nonperformers. This allows us to carry out the tests for the small firms in 8 industries and for the large firms in industries 7 and 8. The results are reported in the last four columns of Table 4. Equality of distributions is rejected in six out of ten cases. Stochastic dominance can hardly be rejected anywhere with the exception of industry 4. Hence, stochastic dominance seems to prevail as long as there are enough observations to meaningfully compare the distributions.

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 eight out of ten cases. This once more makes apparent that omitting R&D expenditures substantially distorts the retrieved productivities.

To obtain a more complete picture of the conditional expectation function \(g(\cdot)\) and its dependence on R&D expenditures, we average this function conditional on three sets of values for \(r_{jt-1}\): When R&D expenditures are zero, when they are above zero but below the median value for performers, and when they are above the median value. That is, for all three sets of values for R&D expenditures we compute

\[
E[g(\omega_{jt-1}, r_{jt-1})|\omega_{jt-1}] = \int g(\omega_{jt-1}, r_{jt-1}) f_{r|\omega}(r_{jt-1}|\omega_{jt-1}) dr_{t-1},
\]
where \( f_{r|\omega}(r_{jt-1}|\omega_{jt-1}) \) is the conditional density of \( r_{jt-1} \) given \( \omega_{jt-1} \) in the sample. The graphs on the left of Figure 2 show in line with intuition that expected productivity is increasing in attained productivity for all three sets of values for R&D expenditures. In addition, with a few exceptions, this relationship shifts upwards with higher R&D expenditures. Sometimes, however, some or part of the dashed lines that represent the expected transitions from current to future productivity for performers lie below the solid line that pertains to nonperformers. This is clearly the case in industries 4, 9, and 10 and once again mirrors the lack of stochastic dominance.

The graphs on the right of Figure 2 depict

\[
E[g(\omega_{jt-1}, r_{jt-1})] = \int g(\omega_{jt-1}, r_{jt-1}) f_{\omega|r}(\omega_{jt-1}|r_{jt-1}) d\omega_{t-1}.
\]

They may be read as the marginal effect of R&D expenditures on expected productivity, once we average out over the values of attained productivity in the sample. Note that the horizontal axis is R&D expenditures in thousands of euros. The dotted line represents the density of R&D expenditure in the sample. The graphs show that higher R&D expenditures tend to be associated with higher expected productivity, even in the case of industry 4 for the bulk of the R&D expenditures. The exceptions here are just industries 9 and 10 with low innovative activity.

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 why expected productivity appears eventually lower in some industries. One possible explanation is heterogeneity across firms within industries, i.e., stochastic dominance may hold if we were able to split these industries into more homogeneous innovative activities.

### 6.4 Productivity growth

Below we assess productivity growth and the role of R&D in it. We compute the expectation of productivity growth as

\[
\beta_t + E(\omega_{jt} - \omega_{jt-1}|\omega_{jt-2}, r_{jt-1}, r_{jt-2}) = \beta_t + E[g(\omega_{jt-1}, r_{jt-1})|\omega_{jt-2}, r_{jt-1}, r_{jt-2}] - g(\omega_{jt-2}, r_{jt-2})
\]

since \( E(\xi_{jt}|\omega_{jt-2}, r_{jt-1}, r_{jt-2}) = E(\xi_{jt-1}|\omega_{jt-2}, r_{jt-1}, r_{jt-2}) = 0 \). For firm \( j \) in period \( t \) this productivity growth is, given attained productivity and R&D expenditures in period \( t - 2 \), a function of R&D expenditures \( r_{jt-1} \) and the innovation to productivity \( \xi_{jt-1} \). By taking the expectation of \( g(\omega_{jt-1}, r_{jt-1}) \) we average out over the realizations of these innovations. We estimate the expectation of productivity growth \( \frac{1}{N} \sum_j \sum_t \frac{1}{T_t} [\hat{g}(\tilde{h}_{jt-1}, r_{jt-1}) - \hat{g}(\tilde{h}_{jt-2}, r_{jt-2})] \). The first three columns of Table 5 report the results for the entire sample and for the subsamples of observations with and without R&D. From hereon we drop 2.5%
of observations at each tail of the distribution.

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 $\mu_{jt} = Y_{jt}/\sum_j Y_{jt}$ are given by the share of output of each firm in each year. Total averages are then computed as $\sum_j \sum_t \frac{1}{T} \mu_{jt} [\hat{g}(\hat{h}_{jt-1}, r_{jt-1}) - \hat{g}(\hat{h}_{jt-2}, r_{jt-2})]$. Columns four to six of Table 5 report the results. The last two columns show a decomposition into the contributions of observations with and without R&D.

The picture that emerges is sensible. Productivity growth, both unweighted and weighted, is higher for performers than for nonperformers in 5 industries, sometimes considerably so. The industries in which this relationship is reversed coincide again with industries 4, 9, and 10 to which we must now add industry 8. 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. More important is the result of the last two columns. The contribution to the expectation of weighted productivity growth of firms that perform R&D is estimated to explain between 70% and 85% of productivity growth in the industries with high innovative activity and between 50% and 65% in the industries with intermediate innovative activity (with the exception of industry 8).

Next we decompose productivity growth to more closely assess the role of R&D. Excluding the trend the growth in expected productivity as given by $g(\omega_{jt-1}, r_{jt-1}) - g(\omega_{jt-2}, r_{jt-2})$ can be decomposed as

$$\left[ g(\omega_{jt-2}, r_{jt-1}) - g(\omega_{jt-2}, r_{jt-2}) \right] + \left[ g(\omega_{jt-1}, r_{jt-1}) - g(\omega_{jt-2}, r_{jt-1}) \right].$$

The first term reflects the change in expected productivity that is attributable to R&D expenditures $r_{jt-1}$, the second the change that is attributable to the “inertia” derived from the previously attained productivity $\omega_{jt-2}$ and the productivity innovation $\xi_{jt-1}$. Table 6 reports unweighted averages of the growth in expected productivity (third column) as well as unweighted averages of the contributions of R&D expenditures (fourth column) and of inertia (fifth column). Column four shows that the contribution of R&D expenditures is positive except for industries 4 and 8. Moreover, the contribution of R&D expenditures represents either a large part of the growth in expected productivity or it may even exceed it and only become smaller because of the negative contribution of inertia.

### 6.5 Rates of return to R&D

Since for any $T > t$ we have $E(\omega_{jT}) = E(\omega_{jT} - \omega_{jT-1}) + E(\omega_{jT-1} - \omega_{jT-2}) + \ldots + E(\omega_{jt} - \omega_{jt-1}) + E(\omega_{jt-1})$, we can view the term $[g(\omega_{jt-2}, r_{jt-1}) - g(\omega_{jt-2}, r_{jt-2})]$ in equation (6) as an increment to productivity that lasts from period $t - 1$ into the indefinite future. Multiplying it by a measure of expected output, say $Y_j$, that is assumed to be constant over time for the sake of simplicity, gives the rent per year that the firm can expect from this amount.
increment to productivity in the absence of depreciation. If the firm optimizes, then the net present value of these rents discounted at a rate $\rho$ and taking into account depreciation at rate $\delta$ must equal R&D expenditures in period $t-1$, i.e.,

$$\sum_{\tau=t-1}^{\infty} \left[ g(\omega_{j\tau-1} - 1, r_{j\tau}) - g(\omega_{j\tau-1}, r_{j\tau-1}) \right] Y_j \left( \frac{1}{1 + \rho + \delta} \right)^{(\tau-t+1)} = R_{t-1}$$

and hence

$$\left[ g(\omega_{j\tau-2} - 2, r_{j\tau-1}) - g(\omega_{j\tau-2}, r_{j\tau-2}) \right] Y_j \frac{R_{t-1}}{R_{t-1}} \approx \rho + \delta.$$  

This suggests that the rate of return to R&D, or dollars obtained on a yearly basis by spending one dollar on R&D, can be roughly estimated by replacing $Y_j$ on the left-hand side of this expression by an appropriate estimate, e.g., current output.

Columns six and seven of Table 6 report weighted averages of the rates of return to R&D computed in this way, where the weights $\mu_{jt} = R_{jt} / \sum_j R_{jt}$ are given by the share of R&D expenditures of each firm in each year. We mostly obtain sensible rates of return to R&D, well within the range of estimates in the literature (see, e.g., Nadiri (1993) or Griliches (2000)). In a few cases, however, they become close to zero and in one case decidedly negative (industry 4). Exploring the reason for these abnormal values, it became clear that there are important differences in the rates of return to R&D belonging to observations that switch from zero R&D to positive R&D or vice versa. When there is a switch we have an extra bonus in some (but not all) industries or an extra malus in industries 4 and 8. Our model thus makes apparent the particular properties of the corner solution of zero R&D expenditures. Extra positive returns could be linked to the sunk costs implied by starting or re-starting innovative activities, but the interpretation of the extra negative returns is more challenging.

To facilitate the comparison with the existing literature, we have estimated the simplest version of the knowledge capital model as given in equation (1) by regressing the first-difference of the log of output on the first-differences of the logs of inputs and the ratio $R_{jt-1}/Y_{jt-1}$ of R&D expenditures to output. The idea is that the estimated coefficient of this ratio can be interpreted as the rate of return to knowledge capital. Since $\varepsilon_{jt} \Delta C_{jt} = C_{jt-1} \frac{\partial Y_j}{\partial C_j} \Delta C_{jt} \approx \frac{\partial Y_j}{\partial C_j} \frac{\Delta C_{jt}}{Y_{jt-1}}$ and $R_{jt-1}$ approximates $\Delta C_{jt}$, the estimated coefficient is $\frac{\partial Y_j}{\partial C_j}$.

As usually happens when running OLS on first-differences of logs, returns to scale and capital coefficients tend to be low. The coefficients on the ratio of R&D expenditures to output are, however, sensible for most of the sectors. The last column of Table 6 presents the rates of return estimated from the knowledge capital model. When these rates are compared with the rates estimated from our model the differences are not overwhelming with the exception, perhaps, of industries 4 and 8. The question is then whether and why our rates of return should be considered more reliable and whether this justifies the extra effort of pursuing the more structural approach.

One difference is that the knowledge capital model yields an average rate of return as the estimated coefficient on the firm- and year-specific ratio of R&D expenditures to output. The equivalent measure in our model is a weighted average of individual rates that
include all the information contained in the firm- and year-specific values of the conditional expectation function \( g(\cdot) \). This reveals a more fundamental difference. Our specification is able to handle and describe heterogeneity. We model at the same time performers and nonperformers, and different amounts of R&D expenditures. We obtain the distribution of productivities in its entirety. This allows us to compute different rates of return for different firms or different groups of firms.

It is worth noting that our individual rates are computed from more reliable coefficient estimates than what the simplest knowledge capital model provides. More generally, our model is more resilient to the possibility of endogeneity bias in assessing the role of R&D. It is also more structural in that we control for the presence of unobserved productivity that likely follows an autocorrelated process. We are thus more sure about the causality of the relationships between expected productivity, current productivity, and R&D expenditures that we find. Of course, 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.

In sum, our approach does not contradict the knowledge capital model. On the contrary, as we have shown in Section 2, a simple extension of the knowledge capital model that allows the accumulation of improvements to productivity to be subjected to shocks is a special case of our model. Our model is richer, in particular with regard to the absence of functional form restrictions and the treatment of heterogeneity. Yet, it is simple enough to apply.

## 7 Concluding remarks

In this paper we develop simple estimators for production functions. The basic idea is to exploit the fact that decisions on variable inputs such as labor and materials are based on current productivity. This results in input demands that are invertible functions and that can thus be used to control for unobserved productivity in the estimation. Moreover, the parametric specification of the production function implies a known form for these functions. This contrasts with the nonparametric methods in the existing literature. The known form of the inverse input demand functions renders identification and estimation more tractable because fewer parts of the model have to now be estimated nonparametrically. As a result, we are able to accommodate a controlled Markov process, thereby capturing the impact of R&D on the evolution of productivity.

We illustrate our approach to production function estimation on an unbalanced panel of Spanish manufacturing firms that cover a total of more than 1800 firms in nine industries during the 1990s. We obtain sensible parameters estimates. Our estimator thus appears to work well. Yet, it is simple to use.

A number of interesting findings emerge. To begin with the R&D process must be treated as inherently uncertain. We estimate that, depending on the industry, between 20% and 50% of the variance in actual productivity is explained by productivity innovations that
cannot be predicted when R&D expenditures are carried out.

Despite this, the distribution of expected productivity of firms that perform R&D tends to stochastically dominate the distribution of expected productivity of firms that do not perform R&D. That is, while the distributions of productivities are overlapping, performers have a higher probability of achieving any particular value of productivity than nonperformers. There is an apparent tendency for the mean of the distribution of expected productivity to be higher for performers than for nonperformers. We estimate this difference in expected productivity to be around 5% in most cases and up to 9% in some cases. There is no clear rule when it comes to the variance. Thus, R&D may not inject additional uncertainty into the evolution of productivity over time.

Expected productivity is increasing in attained productivity and R&D expenditures. Specifically, we find that the transitions from current to future productivity tend to become in expectation more favorable with R&D expenditures. While the conditional expectation function takes a simple separable form in some cases, in most cases the impact of current R&D on future productivity depends crucially on current productivity. Hence, current and past investments in knowledge interact in a complex fashion.

R&D expenditures stimulate productivity growth. The growth in expected productivity corresponding to observations with R&D expenditures is often higher than the growth corresponding to observations without R&D. In addition, we estimate that the contribution of firms that perform R&D explains between 50% and 85% of productivity growth in the industries with intermediate or high innovative activity. Decomposing the expected productivity changes, we are able to separate the impact of carrying out investments in knowledge. Its average is positive and large in almost all industries.

Consequently, we find large average rates of return to R&D in most industries, although our estimates are within the range of the previous literature. Hidden behind these averages, however, is a substantial amount of heterogeneity across firms. Our approach allows us to obtain the distribution of productivities in its entirety. Hence, we are able to compute different rates of return for different firms in different years. In particular, we find that in many industries there is a striking difference between the rates of return for starting and for continuing R&D. The rates of return for starting R&D are higher, possibly due to sunk costs.

Our method can be applied to other contexts. For example, to model and test for two types of technological progress in production functions: Hicks-neutral technological progress that shifts the production function in its entirety and labor-saving technological progress that shifts the ratio of labor to capital. Economists have been for a long time interested in disentangling these effects. In ongoing work we have begun to explore how this can be done by further exploiting the known form of the inverse input demand functions for labor and materials to recover two unobservables, one for Hicks-neutral and one for labor-saving technological progress.
Appendix A

Let \( x_{jt} = (l_{jt}, k_{jt}, w_{jt} - p_{jt}) \) be the vector of arguments of \( h(\cdot) \) and, for the sake of simplicity, surpress R&D expenditures as an argument of \( g(\cdot) \) and \( \tilde{g}(\cdot) \). Then we have \( \omega_{jt} = g(h(x_{jt-1}) + u_{jt}) + \xi_{jt} \) and \( \omega_{jt} = \tilde{g}(h(x_{jt-1})) + \tilde{\xi}_{jt} \). It follows that \( \tilde{\xi}_{jt} = \xi_{jt} + g(h(x_{jt-1}) + u_{jt-1}) - \tilde{g}(h(x_{jt-1})) \) and, taking expectations,

\[
E(\tilde{\xi}_{jt} | x_{jt-k}) = E(\xi_{jt} | x_{jt-k}) + E[g(h(x_{jt-1}) + u_{jt-1}) | x_{jt-k}] - E[\tilde{g}(h(x_{jt-1})) | x_{jt-k}].
\] (7)

Our goal is to show that \( E(\tilde{\xi}_{jt} | x_{jt-k}) \) is zero for all \( k \geq 1 \).

Let \( k = 1 \). The first term in equation (7) is zero because \( \xi_{jt} \) is the error term of the expectation of \( \omega_{jt} \) conditional on \( h(x_{jt-1}) + u_{jt-1} \). But \( E[\omega_{jt} | h(x_{jt-1}) + u_{jt-1}] = E[\omega_{jt} | x_{jt-1}, u_{jt-1}] \) at the true (but unknown) \( h(\cdot) \) by the construction of our Markov process. The second term is \( \tilde{g}(h(x_{jt-1})) \) by equation (4) and thus cancels with the third term. Hence we have \( E(\tilde{\xi}_{jt} | x_{jt-1}) = 0 \).

Let \( k = 2 \). The first term in equation (7) is the expectation of the productivity innovation in period \( t \) and must be orthogonal to all information available in period \( t - 2 \). Using \( f(\cdot) \) to denote a generic density, the second term is

\[
\int g(h(x_{jt-1}) + u_{jt-1}) f_{x,u|x(x_{jt-1}, u_{jt-1} | x_{jt-2})} dx_{jt-1} du_{jt-1}.
\]

Since \( u_{jt-1} \) is by assumption independent from \( x_{jt-1} \) and its lags, it can be written as

\[
\int \left[ \int g(h(x_{jt-1}) + u_{jt-1}) f_{u|x}(u_{jt-1} | x_{jt-2}) du_{jt-1} \right] f_{x|x}(x_{jt-1} | x_{jt-2}) dx_{jt-1} = E[\tilde{g}(h(x_{jt-1})) | x_{jt-2}],
\]

where we use the fact that

\[
\begin{align*}
& f_{x,u|x}(x_{jt-1}, u_{jt-1} | x_{jt-2}) = f_{x,x,u}(x_{jt-1}, x_{jt-2}, u_{jt-1}) f_{x}(x_{jt-2}) \\
& = f_{x,x}(x_{jt-1}, x_{jt-2}) f_{u}(u_{jt-1}) f_{x}(x_{jt-2}) = f_{x,x}(x_{jt-2}) f_{x}(x_{jt-1}) f_{x}(x_{jt-2}) \frac{f_{u}(u_{jt-1}) f_{x}(x_{jt-1})}{f_{x}(x_{jt-2})} \\
& = f_{x,x}(x_{jt-2}) f_{u,x}(u_{jt-1}, x_{jt-1}) f_{x}(x_{jt-2}) f_{x}(x_{jt-1}) = f_{u|x}(u_{jt-1} | x_{jt-1}) f_{x|x}(x_{jt-1} | x_{jt-2}) f_{x}(x_{jt-2}) f_{x}(x_{jt-1}).
\end{align*}
\]

Hence we also have \( E(\tilde{\xi}_{jt} | x_{jt-2}) = 0 \). The extension to higher lags is straightforward.

Appendix B

Table A gives the equivalence between our grouping of industries and the manufacturing breakdown of ESEE. We exclude industry 5 because of data problems.

In what follows we define the variables.

- **Capital.** Capital at current replacement values is computed recursively from an initial estimate and the data on current investments in equipment goods (but not buildings or financial assets), updating the value of past stocks by means of a price index of investment in equipment goods, and using sector-specific estimates of the rates of depreciation. That is, \( K_{jt} = (1 - \delta) \frac{p_{jt}}{p_{jt-1}} K_{jt-1} + I_{jt-1}, \) where capital and investment are in current values and \( p_{jt} \) is the price index of investment. Real capital is obtained by deflating capital at current replacement values by the price index of investment.
- **Investment.** Value of current investments in equipment goods deflated by the price index of investment as needed.

- **Labor.** Number of workers multiplied by hours per worker (normal hours of work plus overtime minus lost hours per worker).

- **Materials.** Value of intermediate consumption deflated by the firm’s price index of materials.

- **Output.** Value of produced goods and services computed as sales plus the variation of inventories deflated by the firm’s price index of output.

- **Price of investment.** Equipment goods component of the index of industry prices computed and published by the Ministry of Industry.

- **Wage.** Hourly wage rate (total labor cost divided by effective total hours of work).

- **Price of materials.** Paasche-type price index computed starting from the percentage variations in the prices of purchased materials, energy, and services as reported by the firm.

- **Price of output.** Paasche-type price index computed starting from the percentage price changes that the firm reports to have made in the markets in which it operates.

- **R&D expenditures.** Total R&D expenditures including the cost of intramural R&D activities, payments for outside R&D contracts, and expenditures on imported technology (patent licenses and technical assistance).

- **Market dynamism.** Weighted index of the dynamism of the markets (slump, stability, and expansion) as reported by the firm for the markets in which it operates.
References


<table>
<thead>
<tr>
<th>Industry</th>
<th>Obs.</th>
<th>Firms</th>
<th>Entry</th>
<th>Exit</th>
<th>Rates of growth</th>
<th>With R&amp;D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Metals and metal products</td>
<td>1235</td>
<td>289</td>
<td>88</td>
<td>17</td>
<td>0.050</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.238)</td>
<td>(0.183)</td>
</tr>
<tr>
<td>2. Non-metallic minerals</td>
<td>670</td>
<td>140</td>
<td>20</td>
<td>15</td>
<td>0.039</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.209)</td>
<td>(0.152)</td>
</tr>
<tr>
<td>3. Chemical products</td>
<td>1218</td>
<td>275</td>
<td>64</td>
<td>15</td>
<td>0.068</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.196)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>4. Agric. and ind. machinery</td>
<td>576</td>
<td>132</td>
<td>36</td>
<td>6</td>
<td>0.059</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.275)</td>
<td>(0.170)</td>
</tr>
<tr>
<td>6. Transport equipment</td>
<td>637</td>
<td>148</td>
<td>39</td>
<td>10</td>
<td>0.087</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0354)</td>
<td>(0.207)</td>
</tr>
<tr>
<td>7. Food, drink and tobacco</td>
<td>1408</td>
<td>304</td>
<td>47</td>
<td>22</td>
<td>0.025</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.224)</td>
<td>(0.186)</td>
</tr>
<tr>
<td>8. Textile, leather and shoes</td>
<td>1278</td>
<td>293</td>
<td>77</td>
<td>49</td>
<td>0.020</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.233)</td>
<td>(0.192)</td>
</tr>
<tr>
<td>9. Timber and furniture</td>
<td>569</td>
<td>138</td>
<td>52</td>
<td>18</td>
<td>0.038</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.278)</td>
<td>(0.210)</td>
</tr>
<tr>
<td>10. Paper and printing products</td>
<td>665</td>
<td>160</td>
<td>42</td>
<td>10</td>
<td>0.035</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.183)</td>
<td>(0.140)</td>
</tr>
</tbody>
</table>
## Table 2
Industry production function estimates (output, 1991-1999)
Using the demand for labor to control for unobserved productivity

<table>
<thead>
<tr>
<th>Industry</th>
<th>No control</th>
<th>Exogenous Markov process</th>
<th>Markov process including R&amp;D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$l$</td>
<td>$k$</td>
<td>$m$</td>
</tr>
<tr>
<td>Metals and metal products</td>
<td>0.251</td>
<td>0.109</td>
<td>0.643</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.013)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Non-metallic minerals</td>
<td>0.277</td>
<td>0.091</td>
<td>0.662</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.020)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Chemical products</td>
<td>0.239</td>
<td>0.060</td>
<td>0.730</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.010)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Agric. and ind. machinery</td>
<td>0.284</td>
<td>0.051</td>
<td>0.671</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.017)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>Transport equipment</td>
<td>0.289</td>
<td>0.080</td>
<td>0.636</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.023)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Food, drink and tobacco</td>
<td>0.173</td>
<td>0.095</td>
<td>0.740</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.014)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Textile, leather and shoes</td>
<td>0.327</td>
<td>0.063</td>
<td>0.607</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.010)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Timber and furniture</td>
<td>0.283</td>
<td>0.079</td>
<td>0.670</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.019)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Paper and printing products</td>
<td>0.321</td>
<td>0.092</td>
<td>0.621</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.016)</td>
<td>(0.025)</td>
</tr>
</tbody>
</table>

1 Estimates for industries 1, 2, 3, 6 and 10 include a time trend.
2 OLS of log of output on a constant, the log of the variables and (for the indicated sectors) a time trend.
3 Reported coefficients are optimal nonlinear GMM estimates and standard errors are robust to heteroskedasticity and autocorrelation.
4 Results for sectors 4, 9 and 10 are from the separable model.
<table>
<thead>
<tr>
<th>Industry</th>
<th>Overidentifying restrictions</th>
<th>Mis specification test</th>
<th>R&amp;D test</th>
<th>Separability test</th>
<th>Var(ω)</th>
<th>Var(e)</th>
<th>Var(ξ) / Var(ω)</th>
<th>Corr(g, ξ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Metals and metal products</td>
<td>106.19 (72)</td>
<td>7.04 0.071</td>
<td>38.21 0.001</td>
<td>3.71 0.295</td>
<td>0.0283</td>
<td>0.0305</td>
<td>0.240 0.325</td>
<td>0.325</td>
</tr>
<tr>
<td>2. Non-metallic minerals</td>
<td>76.29 (72)</td>
<td>1.41 0.702</td>
<td>204.51 0.000</td>
<td>24.52 0.000</td>
<td>0.0306</td>
<td>0.0324</td>
<td>0.324 0.150</td>
<td>0.227</td>
</tr>
<tr>
<td>3. Chemical products</td>
<td>77.13 (72)</td>
<td>2.93 0.403</td>
<td>81.06 0.000</td>
<td>9.88 0.020</td>
<td>0.0224</td>
<td>0.0175</td>
<td>0.219 0.227</td>
<td></td>
</tr>
<tr>
<td>4. Agric. and ind. machinery</td>
<td>79.81 (64)</td>
<td>6.18 0.103</td>
<td>110.81 0.000</td>
<td>0.65 0.885</td>
<td>0.0122</td>
<td>0.0304</td>
<td>0.3930 -0.0961</td>
<td></td>
</tr>
<tr>
<td>6. Transport equipment</td>
<td>81.75 (72)</td>
<td>3.38 0.337</td>
<td>485.44 0.000</td>
<td>211.57 0.000</td>
<td>0.0207</td>
<td>0.0375</td>
<td>0.546 -0.027</td>
<td></td>
</tr>
<tr>
<td>7. Food, drink and tobacco</td>
<td>87.72 (73)</td>
<td>10.61 0.014</td>
<td>49.08 0.000</td>
<td>12.30 0.006</td>
<td>0.0140</td>
<td>0.0255</td>
<td>0.293 0.136</td>
<td></td>
</tr>
<tr>
<td>8. Textile, leather and shoes</td>
<td>90.86 (73)</td>
<td>2.64 0.450</td>
<td>79.69 0.000</td>
<td>21.90 0.000</td>
<td>0.0376</td>
<td>0.0461</td>
<td>0.226 0.457</td>
<td></td>
</tr>
<tr>
<td>9. Timber and furniture</td>
<td>31.78 (29)</td>
<td>2.73 0.436</td>
<td>230.39 0.000</td>
<td>7.54 0.057</td>
<td>0.0136</td>
<td>0.0278</td>
<td>0.4750 0.1094</td>
<td></td>
</tr>
<tr>
<td>10. Paper and printing products</td>
<td>28.76 (28)</td>
<td>15.18 0.002</td>
<td>43.97 0.000</td>
<td>26.57 0.000</td>
<td>0.0270</td>
<td>0.0225</td>
<td>0.4190 0.1628</td>
<td></td>
</tr>
</tbody>
</table>
Table 4
Comparing productivity levels without and with R&D, 1991-1999

<table>
<thead>
<tr>
<th>Industry</th>
<th>Size</th>
<th>Firms with</th>
<th>Diff. of means</th>
<th>Standard dev.</th>
<th>Mean with R&amp;D is greater</th>
<th>Var. with R&amp;D is greater</th>
<th>Kolgomorov-Smirnov tests2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No R&amp;D</td>
<td>R&amp;D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Metals and metal products</td>
<td>≤ 200</td>
<td>143</td>
<td>71</td>
<td>0.045</td>
<td>0.083</td>
<td>0.085</td>
<td>-6.050 1.000 0.943 0.691 1.817 0.003 0.388 0.740</td>
</tr>
<tr>
<td></td>
<td>&gt; 200</td>
<td>11</td>
<td>64</td>
<td>0.038</td>
<td>0.107</td>
<td>0.079</td>
<td>-3.455 1.000 1.835 0.000</td>
</tr>
<tr>
<td>2. Non-metallic minerals</td>
<td>≤ 200</td>
<td>65</td>
<td>27</td>
<td>0.090</td>
<td>0.119</td>
<td>0.105</td>
<td>-6.771 1.000 1.279 0.089 1.705 0.006 0.324 0.811</td>
</tr>
<tr>
<td></td>
<td>&gt; 200</td>
<td>9</td>
<td>39</td>
<td>0.045</td>
<td>0.071</td>
<td>0.070</td>
<td>-4.964 1.000 1.023 0.446</td>
</tr>
<tr>
<td>3. Chemical products</td>
<td>≤ 200</td>
<td>91</td>
<td>81</td>
<td>0.047</td>
<td>0.087</td>
<td>0.076</td>
<td>-7.555 1.000 1.305 0.010 1.673 0.007 0.144 0.959</td>
</tr>
<tr>
<td></td>
<td>&gt; 200</td>
<td>5</td>
<td>98</td>
<td>0.033</td>
<td>0.098</td>
<td>0.095</td>
<td>-2.276 0.987 1.056 0.374</td>
</tr>
<tr>
<td>4. Agric. and ind. machinery</td>
<td>≤ 200</td>
<td>39</td>
<td>56</td>
<td>-0.027</td>
<td>0.104</td>
<td>0.091</td>
<td>2.746 0.003 1.312 0.031 1.085 0.190 1.085 0.095</td>
</tr>
<tr>
<td></td>
<td>&gt; 200</td>
<td>6</td>
<td>31</td>
<td>-0.080</td>
<td>0.083</td>
<td>0.067</td>
<td>5.095 0.000 1.561 0.040</td>
</tr>
<tr>
<td>6. Transport equipment</td>
<td>≤ 200</td>
<td>37</td>
<td>32</td>
<td>0.081</td>
<td>0.090</td>
<td>0.070</td>
<td>-7.915 1.000 1.648 0.005 2.236 0.000 0.000 1.000</td>
</tr>
<tr>
<td></td>
<td>&gt; 200</td>
<td>14</td>
<td>65</td>
<td>0.052</td>
<td>0.077</td>
<td>0.061</td>
<td>-6.135 1.000 1.578 0.002</td>
</tr>
<tr>
<td>7. Food, drink and tobacco</td>
<td>≤ 200</td>
<td>155</td>
<td>49</td>
<td>0.013</td>
<td>0.092</td>
<td>0.074</td>
<td>-1.631 0.947 1.558 0.003 0.960 0.315 0.308 0.827</td>
</tr>
<tr>
<td></td>
<td>&gt; 200</td>
<td>29</td>
<td>71</td>
<td>0.024</td>
<td>0.082</td>
<td>0.060</td>
<td>-3.487 1.000 1.831 0.000 1.012 0.258 0.714 0.361</td>
</tr>
<tr>
<td>8. Textile, leather and shoes</td>
<td>≤ 200</td>
<td>165</td>
<td>59</td>
<td>0.046</td>
<td>0.094</td>
<td>0.119</td>
<td>-4.929 1.000 0.620 1.000 1.762 0.004 0.127 0.968</td>
</tr>
<tr>
<td></td>
<td>&gt; 200</td>
<td>23</td>
<td>46</td>
<td>0.004</td>
<td>0.122</td>
<td>0.104</td>
<td>-0.394 0.653 1.373 0.016 0.681 0.743 0.596 0.492</td>
</tr>
<tr>
<td>9. Timber and furniture</td>
<td>≤ 200</td>
<td>112</td>
<td>18</td>
<td>-0.036</td>
<td>0.082</td>
<td>0.045</td>
<td>4.128 0.000 3.347 0.000</td>
</tr>
<tr>
<td></td>
<td>&gt; 200</td>
<td>1</td>
<td>7</td>
<td>0.004</td>
<td>0.039</td>
<td>0.064</td>
<td>-0.212 0.581 0.371 0.928</td>
</tr>
<tr>
<td>10. Paper and printing products</td>
<td>≤ 200</td>
<td>98</td>
<td>24</td>
<td>0.007</td>
<td>0.104</td>
<td>0.037</td>
<td>-0.962 0.830 8.121 0.000 1.251 0.088 1.023 0.123</td>
</tr>
<tr>
<td></td>
<td>&gt; 200</td>
<td>16</td>
<td>22</td>
<td>-0.018</td>
<td>0.126</td>
<td>0.046</td>
<td>1.345 0.092 7.411 0.000</td>
</tr>
</tbody>
</table>

1 Computed with all observations.
2 Computed with the distribution of firm’s time means when the compared samples have more than 20 firms each.
Table 5
Assesing the role of R&D in expected productivity growth, 1991-1999

<table>
<thead>
<tr>
<th>Industry</th>
<th>Prod. growth(^1)</th>
<th>Prod. growth (weighted)(^2)</th>
<th>Prod. growth (decomp.)(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Metals and metal products</td>
<td>0.0126</td>
<td>0.0183</td>
<td>0.0112</td>
</tr>
<tr>
<td></td>
<td>(0.1340)</td>
<td>(0.1435)</td>
<td>(0.1315)</td>
</tr>
<tr>
<td>2. Non-metallic minerals</td>
<td>0.0185</td>
<td>0.0213</td>
<td>0.0177</td>
</tr>
<tr>
<td></td>
<td>(0.1780)</td>
<td>(0.1670)</td>
<td>(0.1811)</td>
</tr>
<tr>
<td>3. Chemical products</td>
<td>0.0183</td>
<td>0.0232</td>
<td>0.0157</td>
</tr>
<tr>
<td></td>
<td>(0.1488)</td>
<td>(0.1517)</td>
<td>(0.1471)</td>
</tr>
<tr>
<td>4. Agric. and ind. machinery</td>
<td>0.0123</td>
<td>0.0074</td>
<td>0.0164</td>
</tr>
<tr>
<td></td>
<td>(0.1840)</td>
<td>(0.1592)</td>
<td>(0.2021)</td>
</tr>
<tr>
<td>6. Transport equipment</td>
<td>0.0233</td>
<td>0.0347</td>
<td>0.0155</td>
</tr>
<tr>
<td></td>
<td>(0.1475)</td>
<td>(0.1688)</td>
<td>(0.1303)</td>
</tr>
<tr>
<td>7. Food, drink and tobacco</td>
<td>0.0081</td>
<td>0.0102</td>
<td>0.0078</td>
</tr>
<tr>
<td></td>
<td>(0.1425)</td>
<td>(0.1376)</td>
<td>(0.1432)</td>
</tr>
<tr>
<td>8. Textile, leather and shoes</td>
<td>0.0120</td>
<td>0.0030</td>
<td>0.0145</td>
</tr>
<tr>
<td></td>
<td>(0.1434)</td>
<td>(0.1528)</td>
<td>(0.1406)</td>
</tr>
<tr>
<td>9. Timber and furniture</td>
<td>0.0088</td>
<td>0.0052</td>
<td>0.0090</td>
</tr>
<tr>
<td></td>
<td>(0.1287)</td>
<td>(0.1724)</td>
<td>(0.1250)</td>
</tr>
<tr>
<td>10. Paper and printing products</td>
<td>0.0141</td>
<td>0.0136</td>
<td>0.0142</td>
</tr>
<tr>
<td></td>
<td>(0.1829)</td>
<td>(0.1515)</td>
<td>(0.1862)</td>
</tr>
</tbody>
</table>

\(^1\) Unweighted means and standard errors computed with 95\% of the data.

\(^2\) Weighted means of 95\% of the data.
## Table 6
Decomposing productivity growth and computing rates of return to R&D, 1991-1999

<table>
<thead>
<tr>
<th>Industry</th>
<th>Total growth&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Trend</th>
<th>(a^2)</th>
<th>(b^2)</th>
<th>(c^2)</th>
<th>Rates of return&lt;sup&gt;3&lt;/sup&gt;</th>
<th>Knowledge cap. model&lt;sup&gt;5&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R&amp;D obs. (from Table 5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(\frac{g(\omega_{t-1},r_{t-1})-g(\omega_{t-2},r_{t-2})}{R_{jt-1}})</td>
<td></td>
</tr>
<tr>
<td>1. Metals and metal products</td>
<td>0.0183</td>
<td>0.0078</td>
<td>0.0105</td>
<td>0.0036</td>
<td>0.0069</td>
<td>0.3496</td>
<td>0.2166</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Non-metallic minerals</td>
<td>0.0213</td>
<td>0.0187</td>
<td>0.0026</td>
<td>0.0054</td>
<td>-0.0027</td>
<td>0.6372</td>
<td>-0.0291</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Chemical products</td>
<td>0.0232</td>
<td>0.0164</td>
<td>0.0068</td>
<td>0.0018</td>
<td>0.0049</td>
<td>0.1620</td>
<td>0.1631</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Agric. and ind. machinery</td>
<td>0.0074</td>
<td>0.0074</td>
<td>-0.0066</td>
<td>0.0140</td>
<td>-0.1243</td>
<td>0.4724</td>
<td>0.2034</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Transport equipment</td>
<td>0.0347</td>
<td>0.0250</td>
<td>0.0097</td>
<td>0.0137</td>
<td>-0.0040</td>
<td>0.4724</td>
<td>0.2034</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Food, drink and tobacco</td>
<td>0.0102</td>
<td>0.0102</td>
<td>0.0020</td>
<td>0.0083</td>
<td>0.4132</td>
<td>0.183</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Textile, leather and shoes</td>
<td>0.0030</td>
<td>0.0030</td>
<td>-0.0069</td>
<td>0.0099</td>
<td>0.0844</td>
<td>0.3223</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Timber and furniture</td>
<td>0.0052</td>
<td>0.0052</td>
<td>0.0105</td>
<td>-0.0053</td>
<td>0.6906</td>
<td>0.5614</td>
<td>-0.209</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Paper and printing products</td>
<td>0.0136</td>
<td>0.0107</td>
<td>0.0029</td>
<td>0.0102</td>
<td>-0.0073</td>
<td>0.2406</td>
<td>0.0011</td>
</tr>
</tbody>
</table>

<sup>1</sup>Total productivity growth equals trend + \(a\).

<sup>2</sup>Columns \(a\) to \(c\) are unweighted means of the 95% of the data; \(a = b + c\).

<sup>3</sup>Rates of return are averages of the individual rates weighted by the individual R&D expenditures over total R&D expenses.

<sup>4</sup>Continuing observations are the R&D investment observations at \(t\) which follow an investment at \(t - 1\).

<sup>5</sup>OLS using first differences of output, labor, capital and materials. The variable for knowledge capital is \(\frac{R_{jt-1}}{Y_{jt-1}}\).
Figure 1: Expected productivities. Density (left panels) and distribution (right panels).
Figure 1: (cont’d) Expected productivities. Density (left panels) and distribution (right panels).
Figure 1: (cont’d) Expected productivities. Density (left panels) and distribution (right panels).
Figure 2: Conditional expectation function. \( E[g(\omega_{jt-1}, r_{jt-1})|\omega_{jt-1}] \) (left panels) and \( E[g(\omega_{jt-1}, r_{jt-1})|r_{jt-1}] \) (right panels).
Figure 2: (cont'd) Conditional expectation function. $E[g(\omega_{jt-1}, r_{jt-1})|\omega_{jt-1}]$ (left panels) and $E[g(\omega_{jt-1}, r_{jt-1})|r_{jt-1}]$ (right panels).
Figure 2: (cont’d) Conditional expectation function. $E[g(\omega_{jt-1}, r_{jt-1})|\omega_{jt-1}]$ (left panels) and $E[g(\omega_{jt-1}, r_{jt-1})|r_{jt-1}]$ (right panels).