

# Estimating dynamic R&D choice: an analysis of costs and long-run benefits

Bettina Peters\*

Mark J. Roberts\*\*

Van Anh Vuong\*\*\*

and

Helmut Fryges\*\*\*\*

*This article estimates a dynamic structural model of discrete Research and Development (R&D) investment and quantifies its cost and long-run benefit for German manufacturing firms. The model incorporates linkages between R&D choice, product and process innovations, and future productivity and profits. The long-run payoff to R&D is the proportional difference in expected firm value generated by the investment. It increases firm value by 6.7% for the median firm in high-tech industries but only 2.8% in low-tech industries. Simulations show that reductions in maintenance costs of innovation significantly raise investment rates and productivity, whereas reductions in startup costs have little effect.*

## 1. Introduction

■ Firm investment in R&D is a key mechanism generating improvements in firm performance over time. Estimating the *ex post* return to the firm's R&D investment has been a major focus of empirical studies for decades, with most of the empirical literature built around the knowledge

---

\*Centre for European Economic Research (ZEW); b.peters@zew.de.

\*\*Pennsylvania State University and NBER; mroberts@psu.edu.

\*\*\*University of Cologne; vananh.vuong@ewi.uni-koeln.de.

\*\*\*\*University of Tasmania; helmut.fryges@utas.edu.au.

We are grateful to Uli Doraszelski, Jordi Jaumandreu, Ken Judd, Florin Maican, Jacques Mairesse, Matilda Orth, Joris Pinkse, Marc Rysman, Spiro Stefanou, Jim Tybout, Daniel Xu, Hongsong Zhang, and three referees for helpful comments and discussions. We thank participants at the ZEW Conference on Innovation and Patenting (Mannheim, 2013), IIOC (Boston, 2013), EEA (Gothenburg, 2013), German Economic Association (Düsseldorf, 2013), EARIE (Evora, 2013), CDM Workshop (London, 2013), Conference on Innovation and Entrepreneurship (Copenhagen, 2014), MaCCI Annual Conference (Mannheim, 2014), Berlin IO Day (2016), as well as seminars at universities of Cologne (2013), Brussels (2013), Valencia (2013), Turin (2013), Barcelona (2013), Cornell (2013), Leuven (2013), Tinbergen Institute Amsterdam (2013), Heidelberg (2014), Würzburg (2014), Gothenburg (2014), Duke (2014), Oregon (2014), ETH (2016), Zurich (2016), Copenhagen Business School (2016). We thank the Centre for European Economic Research (ZEW) for providing data access and research support.

production function developed by Griliches (1979). In this framework, firm investment in R&D creates a stock of knowledge that enters into the firm's production function as an additional input along with physical capital, labor, and materials. The marginal product of this knowledge input provides a measure of the return to the firm's investment in R&D and has been the primary focus of the empirical innovation literature.

The goal of this article is to estimate the expected payoff to R&D investment at the firm level. Unlike most of the empirical literature that relies on the knowledge production function, we focus on the firm's R&D investment decision. The discrete decision to invest in R&D contains information on the costs of innovation and the expected long-run payoff to the firm from engaging in R&D investment. We develop a dynamic structural model of the firm's choice to invest in R&D, estimate the model using microdata on German manufacturing firms, and summarize the implicit expected long-run payoff to R&D that rationalizes the firm's observed R&D investment decision.

Our model of the firm's dynamic R&D choice captures five important features of the R&D investment process. The first is the impact of R&D on the probability that the firm realizes a product or process innovation. The second is the effect of these realized innovations on the firm's revenue productivity and short-run profitability. Third, these effects can be long lived, affecting the incentives of the firm to invest in the future and impacting the long-run value of the firm. Fourth, there is uncertainty surrounding the effect of R&D on innovation and the effect of innovation on productivity. Fifth, the cost of generating innovations is likely to differ between firms based on their size and whether they are spending to maintain ongoing R&D activities or establishing new R&D programs. Incorporating these features into the model, the structural parameters characterize the linkages between R&D, innovation, and productivity as well as the costs of producing innovations.

We use the model to estimate the long-run payoff to R&D for a sample of German manufacturing firms across a range of high-tech and low-tech industries. The data source is the Mannheim Innovation Panel (MIP) collected by the Centre for European Economic Research (ZEW). This is the German contribution to the Community Innovation Survey (CIS) that is collected for most countries within the Organisation for Economic Co-operation and Development (OECD). The key features of the MIP survey that we utilize are questions on product and process innovations realized by the firm, R&D input measures, production expenditure, capital stocks, and firm sales.

The structural estimates can be briefly summarized. First, firms that invest in R&D have a substantially higher probability of realizing product or process innovations; but R&D investment is neither necessary nor sufficient for firm innovation. The group of high-tech manufacturing industries has a higher probability of innovation, given R&D, than the group of low-tech industries. Second, product as well as process innovations lead to increases in future firm productivity; but product innovations are more important for the high-tech industries, whereas process innovations are more important for the low-tech industries. Third, firm productivity is highly persistent over time, which implies that innovations that raise productivity have long-run effects on firm performance. Fourth, the cost of generating innovations is significantly smaller for firms that are maintaining ongoing R&D investment rather than beginning to invest in R&D. This means that firm R&D history is an important determinant of current R&D behavior.

Using the structural parameters, we estimate the expected payoff to firm R&D as the proportional difference in the expected future value of a firm when it invests in R&D versus when it does not. This expected payoff varies with the productivity, capital stock, age, and industry of the firm and can be constructed for all firms, not just firms that choose to invest. We find that the expected payoff varies substantially across industries and across firms within each industry. For the five high-tech industries, a firm with the median productivity, capital stock, and age has an expected payoff equal to 6.7% of firm value. In the seven low-tech industries, the corresponding payoff is 2.8%. Our results show that the difference between the high-tech and low-tech industries arises from differences in the magnitude of the effect of innovation on the firm's productivity

and profits and differences in the probability of realizing an innovation, given that they invest in R&D.

The estimated dynamic structural model of R&D choice is used to simulate how a change in the cost structure of innovation arising from, for example, a tax break or direct subsidy for R&D investment, affects the firm's investment choice and future productivity growth. We find that, in the high-tech industries, a 20% reduction in the maintenance cost of R&D for firms with R&D experience leads to, after 10 years, an average increase of 9 percentage points in the probability a firm invests in R&D and a 1.4% increase in mean productivity. The same proportional reduction in the innovation cost faced by firms just beginning to invest in R&D has very little impact on the probability of investing or the level of productivity. In the low-tech industries, a reduction of the maintenance cost increases the R&D investment rate by 7 percentage points, but there is little effect on mean productivity. A 20% reduction in the startup cost has a larger effect, raising mean productivity by 2.1% after 5 years and 7.0% after 10 years. The simulations also illustrate that the two changes in cost have very different impacts on firm incentives. Reducing maintenance costs encourages firms to continue or begin investing in R&D. In contrast, the reduction in startup costs encourages new firms to begin investing but also reduces the option value of investing, leading some firms to stop or delay their R&D.

In the next section of the article, we review some key ideas from the empirical literature estimating the private return to R&D. Section 3 develops the theoretical model of R&D investment. Section 4 discusses some important features of the data, and Section 5 develops the empirical model and estimation strategy. Sections 6 and 7 discuss the empirical results and report counterfactual simulations of the model. Section 8 provides concluding remarks.

## 2. The private return to R&D investment

■ The expected private return to a firm's R&D investment is one of the main factors driving the firm's decision to invest. Understanding the magnitude and determinants of the private return is key to explaining the observed patterns of R&D investment and the likely response of firm investment to changes in the economic environment, in particular, policies that subsidize the cost of R&D activities. Estimating the private return to R&D has been a major focus of empirical research for decades, with most of the literature utilizing the knowledge production function framework developed by Griliches (1979).<sup>1</sup> In this framework, firm investment in R&D, or more broadly defined innovation input, creates a stock of knowledge or expertise within the firm that enters into the firm's production function as an additional input along with physical capital, labor, and materials. The key concept of interest in this framework is the partial derivative of output with respect to the knowledge stock, which is estimated as either the elasticity of output with respect to the knowledge stock or the marginal product of the knowledge stock.<sup>2</sup>

This knowledge production function model has been extended in several ways, including incorporating R&D spillovers across firms or industries, using firm market value or Tobin's *q* as a long-run output measure, and incorporating innovation outcomes as an intermediate step between R&D investment and output.<sup>3</sup> With respect to the last extension, the Community Innovation Surveys (CIS) have been developed to collect firm-level information on R&D expenditures and

<sup>1</sup> Surveys of the empirical literature are given by Hall (1996), and Hall, Mairesse, and Mohnen (2010).

<sup>2</sup> The knowledge stock can also depreciate as new products, materials, and production processes make the firm's existing expertise irrelevant. The marginal product can be interpreted as the gross rate of return to R&D, whereas the net rate of return is defined as the marginal product minus the rate of depreciation.

<sup>3</sup> See Griliches (1992) for a discussion of spillovers and Hall, Mairesse, and Mohnen (2010) for a recent review of the empirical evidence. Cohen (2010) provides a broad-ranging review of the empirical literature on innovative activity and performance, including numerous studies linking R&D investment with underlying firm characteristics. Czarnitzki, Hall, and Oriani (2006) review the literature that measures the effect of the knowledge capital stock on firm market value. A large empirical literature has been built around the model of Crépon, Duguet, and Mairesse (1998) to incorporate innovation outcomes into the R&D-productivity process. This literature is reviewed in Hall (2011) and Mohnen and Hall (2013).

innovations, including the development of new products and the adoption of new or improved production processes.<sup>4</sup> Even with these extensions, the primary focus of the literature remains the estimation of either the output elasticity or marginal product of the knowledge capital stock.

An alternative approach to incorporating R&D in the firm's production process has been implemented by Aw, Roberts, and Xu (2011) and Doraszelski and Jaumandreu (2013). They model the firm's productivity as a Markov process that is altered by the firm's endogenous decision to invest in R&D. Conceptually, the stochastic productivity and knowledge capital models are very similar, they explain current output as a function of past R&D investment and past productivity shocks. There are, however, two main advantages of the stochastic productivity approach in the implementation. The basic form of the knowledge capital model treats the stock of knowledge as the undepreciated sum of past R&D expenditures. In practice, it is difficult to precisely estimate distributed lag coefficients on past R&D expenditures, and researchers generally assume a fixed depreciation rate which imposes a lag structure on past R&D. Additionally, it is difficult to estimate the stock of knowledge in the initial time period the firm is observed, and this leads to measurement error in this variable, which may be particularly important when the data contains only short time series of expenditures, as is the case in this application. In contrast, the stochastic productivity framework models current productivity as a function of prior-period productivity, the current period investment in R&D, and a stochastic shock. The effect of distant past R&D expenditures is captured in prior-period productivity and the level of productivity and the persistence in the productivity process can be estimated even with short time series of firm-level data. A second advantage of the stochastic framework is that it allows random shocks to current productivity to carry over into future productivity. This captures an element of the uncertainty in the productivity process and allows for two firms with the same past R&D expenditure path to have different productivity levels.<sup>5</sup> Doraszelski and Jaumandreu (2013) implement hypothesis tests that allow them to discriminate between different variations of the knowledge capital and stochastic productivity specifications. Their results favor the endogenous stochastic productivity model. In this article, we incorporate a stochastic productivity process as one component of the dynamic model of R&D choice.

Griliches (1979) raises concerns about the ability of the knowledge production function model to clarify the simultaneity between output and R&D expenditure. Current R&D expenditures increase the future knowledge stock, which then increases future output through the production function. That is the mechanism of interest, but estimating it is complicated by the fact that current R&D expenditures are determined by past output and the firm's expectation of future output. Griliches warns that without careful attention to model specification and formulation, estimates of the effect of R&D on output in this framework may largely reflect the effect of output on R&D (Griliches, 1979). In this article, we develop an alternative approach that deals with the simultaneity issues by modelling and estimating the firm's dynamic decision to invest in R&D rather than just the production function.<sup>6</sup> In doing so, we model the simultaneous and intertemporal linkages between R&D and output that are identified by Griliches.

---

<sup>4</sup> See Hall (2011) for a survey of the empirical studies and Mairesse and Mohnen (2010) and Mairesse, Mohnen, and Kremp (2005) for a discussion of the estimation issues that arise in using the CIS data. Roberts and Vuong (2013) provide a comparison of the structural model of R&D investment that we develop in this article and the framework from Crépon, Duguet, and Mairesse (1998).

<sup>5</sup> Klette (1996), Griliches (1998), and Rogers (2010) extend the basic knowledge capital model by also incorporating an exogenous stochastic process for productivity. Doraszelski and Jaumandreu (2013) discuss the differences between the two approaches.

<sup>6</sup> Aw, Roberts, and Xu (2011) estimate a dynamic demand curve for R&D by Taiwanese manufacturers. Xu (2008) estimates a dynamic demand curve which includes a private return to R&D and an across-firm spillover that generates potential social benefits from R&D. Hashmi and Van Biesebroeck (2016) estimate a model of dynamic R&D competition among firms in the automobile industry. Peters, Roberts, and Vuong (2017) estimate a dynamic model of R&D demand that allows for differences in financial resources across firms. Bernstein and Nadiri (1989, 1991) estimate a demand curve for R&D using a dynamic cost function model. Their model of R&D investment is analogous to an investment model for physical capital, and they estimate an Euler equation for the knowledge capital stock.

An important contribution of our dynamic framework is that it leads to a more comprehensive measure of the benefit of R&D. Rather than defining it as the marginal product of knowledge capital in the production function, we define it as the impact of the firm's R&D choice on the expected future value of the firm. This benefit depends on how R&D affects productivity and output in the subsequent period, which is the focus of the knowledge production function literature, but also on how the change in productivity impacts the discounted sum of future firm profits, including its effect on the firm's incentives to invest in R&D in the future.

### 3. Theoretical model

■ This section develops a theoretical model of a firm's dynamic decision to undertake R&D investment. In the model, the firm's current productivity is a state variable that impacts the firm's decision to invest in R&D and then evolves endogenously as a result of that decision. The specific mechanism we model is that the firm's decision to invest in R&D alters the probability of the firm realizing a product or process innovation in the future. If the firm realizes an innovation, this shifts the distribution of future productivity and, ultimately, the future profits they earn. The firm chooses to invest in R&D if the expected long-run payoff resulting from this R&D-innovation-productivity process is greater than the current investment cost. This is analogous to the firm deciding in each period whether or not to exercise the option of buying a possible productivity improvement. The expected long-run payoff of the investment reflects the firm's valuation of the productivity improvement and will vary across firms with differences in their characteristics, in particular their capital stock, productivity level, age, and industry affiliation. The cost of generating innovations also varies across firms with differences in their size and prior R&D experience.

The model contains four structural components. The first is the firm's profit function linking productivity and profits,  $\pi(\omega_{it})$ , where  $\omega_{it}$  is firm  $i$ 's productivity in year  $t$ . The second links the firm's R&D decision with the probability it realizes either a product or process innovation in the future. This component is represented by a cdf  $F(d_{it+1}, z_{it+1} | r d_{it})$ , where  $d, z, rd$  are measures of product innovation, process innovation, and R&D choice, respectively. This specification captures the uncertainty surrounding whether or not the firm's R&D efforts will produce innovation. The third component describes the process of productivity evolution, in which product and process innovations affect the probability distribution of the firm's future productivity,  $G(\omega_{it+1} | \omega_{it}, d_{it+1}, z_{it+1})$ .<sup>7</sup> This component leads to variability and uncertainty in the firm's future profits. The final structural component is the cost function for innovation,  $C(r d_{it-1})$ . This cost is either a sunk startup cost or a fixed maintenance cost, depending on the firm's prior history of R&D participation. The next subsections discuss each of these components in more detail.

□ **Productivity and the firm's short-run profits.** The firm's short-run marginal cost is given by

$$c_{it} = \beta_0 + \beta_k k_{it} + \beta_a a_{it} + \beta_w w_t - \psi_{it}, \quad (1)$$

where  $c_{it}$  is the log of marginal cost,  $k_{it}$  is the log of firm capital stock,  $a_{it}$  is firm age, and  $w_t$  is a vector of market prices for variable inputs that every firm faces in period  $t$ . The firm-specific production efficiency  $\psi_{it}$  captures differences in technology or managerial ability and is known by the firm but not observable to the econometrician.<sup>8</sup> The capital stock is treated as a fixed factor

<sup>7</sup> Olley and Pakes (1996) and the empirical applications that follow from them assume an exogenous Markov process for firm productivity,  $G(\omega_{it+1} | \omega_{it})$ . Doraszelski and Jaumandreu (2013) endogenize productivity evolution by including a measure of R&D investment,  $G(\omega_{it+1} | \omega_{it}, r d_{it})$ . Aw, Roberts, and Xu (2011) let productivity evolution depend on the firm's R&D and export market participation,  $G(\omega_{it+1} | \omega_{it}, r d_{it}, ex_{it})$ .

<sup>8</sup> Variation in input quality, which leads to variation in input prices, across firms is also captured in  $\psi$ . We model this source of quality variation as part of the unobserved firm efficiency.

in the short run. Thus, there are three sources of cost heterogeneity across firms: capital stock, age, and unobserved production efficiency.<sup>9</sup>

The demand for firm  $i$ 's product  $q_{it}$  is given by

$$q_{it} = Q_t \left( \frac{p_{it}}{P_t} \right)^\eta \exp(\phi_{it}) = \Phi_t (p_{it})^\eta \exp(\phi_{it}), \quad (2)$$

where  $Q_t$  is the aggregate industry output in period  $t$  and  $P_t$  is the industry price index, which are combined into the industry aggregate  $\Phi_t$ . The firm-specific variables are the firm's output price  $p_{it}$  and a demand shifter  $\phi_{it}$  that reflects product desirability or quality. The demand shifter is known by the firm but not observed by the econometrician. The elasticity of demand  $\eta$  is negative and assumed to be constant for all firms in the industry.

Assuming the firm operates in a monopolistically competitive market, it maximizes its short-run profit by setting the price for its output equal to a constant markup over marginal cost:  $p_{it} = \left( \frac{\eta}{1+\eta} \right) \exp(c_{it})$ . Given this optimal price, the log of the firm's revenue  $r_{it}$  is:

$$r_{it} = (1 + \eta) \ln \left( \frac{\eta}{1 + \eta} \right) + \ln \Phi_t + (1 + \eta)(\beta_0 + \beta_k k_{it} + \beta_a a_{it} + \beta_w w_t - \omega_{it}). \quad (3)$$

Revenue productivity is denoted by  $\omega_{it}$  and is defined as  $\omega_{it} = \psi_{it} - \frac{1}{1+\eta} \phi_{it}$ . Equation (3) implies that, for a given capital stock and age, heterogeneity in the firm's revenue is driven by differences in production efficiency  $\psi$  and the demand shifter  $\phi$ . We refer to the unobserved revenue productivity  $\omega_{it}$  simply as productivity.<sup>10</sup> Given the form of the firm's pricing rule, there is a simple relationship between the firm's short-run profits and revenue:

$$\pi_{it} = \pi(\omega_{it}) = -\frac{1}{\eta} \exp(r_{it}). \quad (4)$$

The link between productivity  $\omega$  and short-run profits is an important determinant of the firm's decision to invest in R&D.

□ **R&D investment and endogenous productivity.** A key component of our framework is that the firm can endogenously affect the evolution of productivity and profits over time by choosing whether or not to participate in R&D activities. We model this linkage in two components. First, the firm makes a discrete decision to invest in R&D,  $rd_{it} \in \{0, 1\}$ , and this affects the probability the firm realizes a product or process innovation. Innovations are denoted as  $z_{it+1}$  and  $d_{it+1}$ , which are discrete variables equal to 1 if firm  $i$  realizes a process or product innovation in year  $t + 1$  and 0 otherwise. The linkage between R&D and innovation is represented by the cumulative joint distribution of product and process innovations, conditional on whether or not the firm invests in R&D,  $F(d_{it+1}, z_{it+1} | rd_{it})$ . We expect that firms that invest in R&D will be more likely to realize product and process innovations in the next period.

We choose to treat the firm's R&D decision as a discrete choice for several reasons. In our data, there are substantial differences in the probabilities of product and process innovation between firms that invest in R&D and firms that do not (evidence is provided in Table 3), but

<sup>9</sup> Equation (1) implies that, in the short run, the firm can expand or contract output at constant marginal cost. This is a reasonable assumption if, along with the variable inputs, the firm can also adjust the utilization of its fixed capital stock in order to expand or contract its output in the short run. In addition, in micropanel data of the type we utilize, most of the variation in firm sales is in the across-firm rather than within-firm dimension. To account for this, our marginal cost model relies on three factors, the capital stock, age, and production efficiency that primarily vary across firms. Economies or diseconomies of scale are unlikely to be the source of the firm sales variation we observe in the data.

<sup>10</sup> Empirical measures of revenue productivity are also likely to reflect differences in markups, which are not included in the theoretical model. To estimate our model of R&D demand, we only need to quantify the effect of  $\omega_{it}$  on firm sales and profit and do not need to separate it into components reflecting  $\psi$  and  $\phi$  or markups. Studies that identify cost and demand shocks using quantity and price data are Foster, Haltiwanger, and Syverson (2008) and Roberts, Xu, Fan, and Zhang (2012).

fluctuation in the level of R&D spending has little effect on these probabilities. This indicates that there are basically two innovation regimes in the data, one for firms that invest in R&D and the other for firms that do not. As R&D in our framework works through the probability of an innovation, it is important to capture the difference between firms that invest in R&D and firms that do not, but there is little additional gain from treating R&D as a continuous variable. Furthermore, measurement error in the level of R&D expenditure is more substantial than the error in the discrete participation variable, and this can result in small estimated impacts of R&D expenditure on the probability of innovation.<sup>11</sup> Overall, the discrete R&D participation variable is a robust indicator of the firm's investment strategy and clearly distinguishes the firms that choose to invest in uncertain R&D projects from those that do not. The focus of our empirical model is on measuring the long-run difference in firm value between firms that adopt these different investment strategies.

This specification of  $F(d_{it+1}, z_{it+1}|rd_{it})$  also recognizes that firms may direct their R&D activity in different ways, including improving their production processes and developing new or improved products. For example, in industries with extensive product differentiation across firms, R&D may be heavily focused on new product development, whereas in large-scale homogeneous-product industries, R&D may be focused on reducing cost through process innovations. This specification also captures one aspect of the uncertainty that firms face when investing in R&D: the technological uncertainty surrounding the innovation process. The cdf must be general enough to recognize that R&D investment is neither necessary nor sufficient for innovation. A firm with R&D investment might not realize any product or process innovations, whereas another firm may realize one or both innovations, even without R&D investment. The latter can result from luck, the effect of expenditures on R&D in the more distant past even if the firm is not currently investing, ideas that are brought to the firm by hiring experienced workers or other spillover channels, or changes in the production process that result from learning-by-doing without formal R&D investment.

The second component of the R&D-productivity linkage is modelled with the cdf  $G(\omega_{it+1}|\omega_{it}, d_{it+1}, z_{it+1})$ , where firm productivity is a stochastic variable that is affected by the firm's past productivity and the current realizations of product and process innovations. This formulation captures a second aspect of the uncertainty that firms face when investing in R&D: uncertainty surrounding the economic value of an innovation. Even when they realize an innovation, the exact impact of that innovation on future productivity and profits is unknown. It may also be the case that product and process innovations have different impacts on future productivity because each works through different channels on the demand and cost sides. We assume that firm productivity evolves as:

$$\omega_{it+1} = g(\omega_{it}, d_{it+1}, z_{it+1}) + \varepsilon_{it+1}. \quad (5)$$

The function  $g(\cdot)$  is the conditional expectation of future productivity and  $\varepsilon$  is a zero mean stochastic shock. This captures several important aspects of productivity evolution. First, the firm's productivity is assumed to persist over time. This intertemporal persistence is an important feature of firm-level data on productivity. Second, innovations are allowed to systematically shift the mean of the distribution of future firm productivity. Expected future productivity evolves only in those cases in which the firm realizes a product or process innovation, capturing the fact that R&D expenditures alone are not sufficient to generate productivity improvements. Third, the specification recognizes that productivity change is affected by stochastic shocks  $\varepsilon_{it+1}$ , which reflect the inherent randomness in the productivity process. We assume that the productivity shocks  $\varepsilon_{it+1}$  are *i.i.d.* across time and firms and are drawn from a normal distribution with zero mean and variance  $\sigma_\varepsilon^2$ . Because of the persistence in productivity, the shocks in any period are incorporated into future productivity levels rather than having transitory effects.

<sup>11</sup> See Mairesse, Mohnen, and Kremp (2005) for a discussion and evidence on this point using firm data from the French innovation survey.

Combining these two stages captures the uncertainty and the endogeneity of the productivity process. By making investments in R&D, the firm alters the probability of receiving a product or process innovation, which in turn alters the distribution of productivity that it faces in future periods. We refer to the first step as the innovation process and the second step as the productivity evolution process. By including the innovation process in the model, rather than linking R&D directly to productivity as in Aw, Roberts, and Xu (2011) and Doraszelski and Jaumandreu (2013), we can gain some additional insight into whether R&D is working to improve productivity through the demand side or cost side of the firm's operations. In this framework, productivity improves with either cost reductions or demand expansions that lead to higher firm sales. Although we cannot measure the impact of R&D on the separate demand and cost components,  $\phi_{it}$  and  $\psi_{it}$ , if we find that the overall linkage between R&D and productivity is primarily due to product innovations, it suggests that R&D is working through the demand side, whereas a finding of a more important role for process innovations suggests R&D is working through the cost side.

□ **The firm's dynamic decision to invest in R&D.** This subsection develops the firm's decision rule for whether or not to invest in R&D. The benefits of investing depend on the effect of R&D on the firm's expected future productivity and profits, as developed in the last two subsections. The firm's decision also depends on the cost of improving its productivity. In this model, the firm's cost is the expenditure it must make to generate a productivity improvement. This cost varies across firms for many reasons, such as the nature of the project or number of projects the firm invests in, the firm's expertise in the innovation process, its possibilities to access financial resources, and its prior R&D experience. The fact that some firms have higher innovation capabilities or have a larger set of technological opportunities for innovation is captured in this model by lower innovation costs. Similarly, it is likely that a firm that performs R&D continuously over time is able to generate an innovation with lower expenditures than a firm that is just beginning to invest in R&D, because it can rely on past expertise or synergy effects from previous projects. To capture heterogeneity in innovation costs, we model a firm's innovation cost  $C_{it}$  as a draw from an exponential distribution where the mean of the distribution varies with the firm's size, measured by its capital stock, and its prior R&D experience. Defining the discrete indicator variable  $rd_{it-1}$  to equal one if the firm invested in R&D in year  $t - 1$  and zero if it did not, the innovation cost  $C_{it}$  of firm  $i$  in year  $t$  can be represented by:

$$C_{it} \sim \exp(\gamma^m * rd_{it-1} * k_{it} + \gamma^s * (1 - rd_{it-1}) * k_{it}). \quad (6)$$

A firm with prior R&D experience has to pay a maintenance cost drawn from a distribution with a mean of  $\gamma^m k_{it}$  and a firm with no prior experience has to pay a startup cost drawn from a distribution with a mean of  $\gamma^s k_{it}$ . The parameter vector  $\gamma = (\gamma^m, \gamma^s)$  captures differences in the maintenance and startup cost distributions. This innovation cost is observed by the firm prior to their investment decision and acts as an additional source of unobserved heterogeneity (to us) that drives the firm's investment behavior.

We assume that, at the start of period  $t$ , the firm observes its current productivity level  $\omega_{it}$ , knows its short-run profit function, and the processes for innovation and productivity evolution  $F$  and  $G$ . The firm's state variables  $s_{it} = (\omega_{it}, rd_{it-1})$  evolve endogenously as the firm makes its decision to conduct R&D,  $rd_{it} \in \{0, 1\}$ .<sup>12</sup> Given its state vector and discount factor  $\beta$ , the firm's value function  $V(s_{it})$ , before it observes the maintenance or startup cost, can be written as:

$$V(s_{it}) = \pi(\omega_{it}) + \int_{C_{it}} \max_{rd \in \{0,1\}} (\beta E_t V(s_{it+1} | \omega_{it}, rd_{it} = 1) - C_{it}; \beta E_t V(s_{it+1} | \omega_{it}, rd_{it} = 0)) dC, \quad (7)$$

<sup>12</sup> Each firm is characterized by three exogenous variables: age  $a_{it}$ , which enters the profit function, capital stock  $k_{it}$ , which enters the profit function and innovation cost function, and industry, which enters all of the structural components. To simplify the notation, we suppress these exogenous characteristics and explain the dynamic decision to invest in R&D, focusing on the endogenous variables in the model,  $\omega$  and  $rd$ . In the empirical model, we treat the firm's capital stock, age, and industry as defining an exogenous firm type and solve the firm's value function  $V(s_{it})$  for each firm type.



where the expected future value of the firm is defined as an expectation over the future levels of productivity and innovation outcomes:

$$E_t V(s_{it+1}|\omega_{it}, rd_{it}) = \sum_{(d,z)} \int_{\omega} V(s_{it+1}) dG(\omega_{it+1}|\omega_{it}, d_{it+1}, z_{it+1}) dF(d_{it+1}, z_{it+1}|rd_{it}). \quad (8)$$

Equation (7) shows that the firm will choose to invest in R&D if the discounted expected future profits from investing,  $\beta E_t V(s_{it+1}|\omega_{it}, rd_{it} = 1)$ , net of the relevant maintenance or startup cost, are greater than the expected future profits from not investing,  $\beta E_t V(s_{it+1}|\omega_{it}, rd_{it} = 0)$ . What differentiates these two expected future profits is the effect of R&D on the firm's future productivity. Using this specification, we can define the marginal benefit of conducting R&D as:

$$\Delta EV(\omega_{it}) \equiv \beta E_t V(s_{it+1}|\omega_{it}, rd_{it} = 1) - \beta E_t V(s_{it+1}|\omega_{it}, rd_{it} = 0). \quad (9)$$

The firm will choose to invest in R&D if  $\Delta EV(\omega_{it}) \geq C_{it}(rd_{it-1})$ . This is the condition used in the empirical model to explain the firm's observed R&D choice.

Overall, this model endogenizes the firm's choice to undertake R&D investments as a comparison between the net expected future profits of the two alternatives. Using the empirical model we develop in Section 5, we estimate the innovation function, productivity evolution process, and distributions of startup and maintenance costs of innovation faced by the firm, and quantify  $\Delta EV(\omega_{it})$ , the expected long-run payoff to investing in R&D.

## 4. Data

□ **Firm sample.** The data we use to analyze the role of R&D in the productivity evolution of German firms are contained in the Mannheim Innovation Panel (MIP) survey collected by the Centre for European Economic Research (ZEW) on behalf of the German Federal Ministry of Education and Research. The survey is conducted every year for firms in the manufacturing, mining, energy, water, construction, and service sectors. Samples are drawn from the Creditreform database according to the stratifying variables firm size, region, and industry.<sup>13</sup> These are representative of firms with German headquarters and at least five employees.

The manufacturing survey begins in 1993 and contributes to the Community Innovation Surveys (CIS) that are administered in many OECD countries. The survey adheres to the *Oslo Manual*, which provides guidelines for the definition, classification, and measurement of innovation (OECD, 1992; 1997; 2005). Every year, the same set of firms is asked to participate in the survey and to complete the questionnaire sent to them via mail. The sample is updated every two years to account for exiting firms, newly founded firms, and firms that developed to satisfy the selection criteria of the sample. Participation in the survey is voluntary and the average response rate is approximately 25%, resulting in approximately 5000 survey responses across all industries in each year. A nonresponse analysis is performed via phone to check and correct for nonresponse bias. Due to cost reasons, starting in 1998, the full questionnaire was only sent out every other year to all firms in the full sample. However, information on variables of interest, such as sales, capital stock, and variable costs, are asked retrospectively for the previous year to ensure the annual coverage. In odd years, only short questionnaires with core questions are sent to a subset of firms. Therefore, the number of firms in odd years in the panel is significantly lower than in even years. This limits the ability to follow individual firms over long periods of time.<sup>14</sup>

For the empirical analysis, we focus on two groups of manufacturing industries. The high-tech (HT) industry group consists of firms in five aggregated two-digit manufacturing industries (NACE codes<sup>15</sup>): chemicals (23, 24), nonelectrical machinery (29), electrical machinery (30, 31,

<sup>13</sup> The Creditreform database is the largest credit rating agency in Germany and maintains a comprehensive database of approximately 3.3 million German firms.

<sup>14</sup> See Rammer and Peters (2013) for further details and summary statistics from the MIP.

<sup>15</sup> The industry definition is based on the classification system NACE Rev. 1 (Nomenclature statistique des Activités économiques dans la Communauté Européenne) as published by Eurostat (1996).

32), instruments (33), and motor vehicles (34, 35). Based on OECD data, these industries all have R&D-sales ratios that exceed 0.025. The low-tech (LT) industry group includes firms in seven aggregated industries, food (15, 16), textiles (17, 18, 19), paper (20, 21, 22), plastic (25), nonmetallic minerals (26), basic metals (27, 28), and miscellaneous manufacturing (36, 37), that all have much lower R&D-sales ratios. We estimate the model separately for the high-tech and low-tech groups and include dummy variables for the aggregated two-digit industries in the profit function and innovation cost function. We also estimate the innovation probabilities separately for each two-digit industry. Our final sample consists of observations between 1993–2008 on all firms in these industries with at least two consecutive observations and nonmissing information on the needed variables. There are a total of 3313 firm-year observations used for estimation in the high-tech industries, 4290 observations for the low-tech industries, and an average of 2.6 observations for each firm.

The sample attrition that occurs is virtually all due to nonreporting and not due to firm death. Beginning in 1999, we use codes in the Creditreform data set and the MIP questionnaire to identify firms that disappear from our sample and are likely to be true firm deaths. Depending on the stringency of our death criteria, we find that only between 1.77% and 5.20% of the observations that disappear from our sample are true or likely firm deaths. We also find that, comparing the firms that remain in the sample and those that exit the sample, there is no significant difference in firm characteristics, particularly productivity, in their last year of observation. The sample attrition in our data set is random and is not due to the death of low-productivity firms and does not generate selection bias in the estimates of the revenue function.

□ **Variable measurement.** For the estimation, we use data on firm revenue, variable costs, capital stock, age, innovation expenditures, and product and process innovations.<sup>16</sup> Firm revenue is the sum of domestic and export sales. Total variable cost is defined as the sum of expenditure on labor, materials, and energy, and the firm's short-run profit is the difference between revenue and total variable cost. The firm's value is the discounted sum of the future short-run profits and thus measures the long-run resources available to pay its capital expenses plus the economic profits. Firm age is measured using a set of four dummy variables distinguishing the age groups: 1–9 years, 10–19 years, 20–49 years, and  $\geq 50$  years.

A special feature of the Community Innovation Surveys is that they provide measures of innovation input and innovation outputs. Innovation input is measured by the firm's expenditure on a set of activities related to innovation. This measure includes R&D spending but also spending on worker training, acquisition of external knowledge and capital, marketing, and design expenditures for producing a new product or introducing a new production process. The R&D variable we analyze in the empirical model ( $rd_{it}$ ) takes the value one if the firm reports a positive level of spending on innovation activities.

Innovation output captures the introduction of a new product or a new production process by the firm.<sup>17</sup> The *Oslo Manual* defines a product innovation as a new or significantly improved product or service. A process innovation refers to new or significant changes in the way products are produced, delivered, or supplied. The main purpose of a process innovation is to reduce production costs or to improve the quality of a product. For instance, the use of lasers to increase the quality of products in metal processing or the introduction of automation concepts is process innovations. The innovation does not have to be new to the market but only to the firm. A firm could report an innovation if it adopted a production technology from a competitor or expanded its product line, even if the product was already offered by other firms.

<sup>16</sup> For 1999 and 2000, the panel does not contain information on the firms' capital stock. We impute these missing years using linear interpolation.

<sup>17</sup> Beginning in 2005, the survey started to also include questions on organizational innovation, which is defined as new business practices, workplace organization, or external relations, and marketing innovation, referring to changes in product design, packaging, product placement or promotion, and pricing methods. The time-series information on these variables is too short for them to be utilized in this study.

**TABLE 1** Innovation Rates By Industry—pooled over Firms and Years

	Proportion with Expenditure $rd$	Proportion with Product Innovation $d$	Proportion with Process Innovation $z$
High-Tech Industries			
Chemicals	0.769	0.686	0.552
Machinery	0.757	0.692	0.524
Electronics	0.826	0.765	0.560
Instruments	0.828	0.783	0.534
Vehicles	0.714	0.621	0.534
<b>Average HT Industries</b>	<b>0.780</b>	<b>0.713</b>	<b>0.539</b>
Low-Tech Industries			
Food	0.536	0.481	0.409
Textiles	0.476	0.416	0.311
Paper	0.480	0.355	0.386
Plastic	0.634	0.591	0.486
Minerals	0.580	0.536	0.452
Basic metals	0.582	0.469	0.472
Misc. manuf.	0.632	0.555	0.408
<b>Average LT Industries</b>	<b>0.560</b>	<b>0.477</b>	<b>0.428</b>

The timing assumptions in the theoretical model about the relationship between R&D spending, innovation outcomes, and productivity are fairly general: R&D spending precedes innovation outcomes, and innovations that are realized are assumed to affect productivity and profits in the period they are introduced. In the survey in year  $t$ , the firms are asked whether they introduced new or significantly improved products or services during the years  $(t - 2)$ ,  $(t - 1)$ , or  $t$ . The discrete variable *product innovation*  $d_{it}$  takes the value one if the firm reports yes to the question. The discrete variable for *process innovation*  $z_{it}$  equals one if the firm reports new or significantly improved internal processes during the years  $(t - 2)$  to  $t$ . In the empirical model, this outcome is related to R&D spending in the previous year  $(t - 1)$ , so there is not a perfect match between the timing of the R&D and the realization of the innovations. This may lead us to overestimate the effect of R&D on innovation, because the innovation variable could be capturing an outcome from one year before the R&D investment was made. Attempting to use more distant lags of R&D spending exaggerates the problems caused by sample attrition and reduces the number of observations containing the necessary current and lagged variables.

Table 1 summarizes the proportion of firms in the sample that report positive innovation expenditures, successful product innovations, and successful process innovations for each industry. The industries are aggregated into the high-tech and low-tech groups. In our sample, the majority of firms report having expenditures on innovation activities, but the proportions differ across industries. In the five high-tech industries, the proportion varies from 0.714 to 0.828, whereas, in the seven low-tech industries, it varies from 0.476 to 0.634. The rate of product innovation is also higher in the high-tech industries: between 0.621 and 0.783 of the firm-year observations report having a new product innovation, whereas, in the low-tech group, the rate of product innovation varies from 0.355 to 0.591. This same difference exists for process innovation, but the difference in magnitude between the high-tech and low-tech industries is not as large. The high-tech industries vary in a narrow band between 0.524 and 0.560, and all but one of the low-tech industries vary between 0.386 and 0.486. The model developed in the last section allows product and process innovations to occur at different rates, given the firm's R&D choice, and allows them to each have a different impact on future productivity. This leads to differences in the expected benefits of R&D across industries and helps to explain differences in the proportion of firms that choose to invest in R&D.

Table 2 summarizes the difference in firm revenue, capital stock, and age by industry and R&D status. The table reports the median values in each category. Differences in capital stock and

TABLE 2 Firm Characteristics—Median over All Observations in the Category

	Capital Stock*		Age		Revenue <sup>a</sup>	
	<i>rd</i> = 1	<i>rd</i> = 0	<i>rd</i> = 1	<i>rd</i> = 0	<i>rd</i> = 1	<i>rd</i> = 0
High-Tech Industries						
Chemicals	8.181	4.474	23	33	32.000	30.768
Machinery	4.136	1.000	19	14	19.648	3.579
Electronics	2.272	0.271	16	14	12.871	3.337
Instruments	1.636	0.358	18	15	9.985	2.127
Vehicles	9.383	1.170	18	17	52.034	4.346
Low-Tech Industries						
Food	5.113	2.700	20	16	15.200	9.076
Textiles	1.855	0.794	16	39	7.719	7.490
Paper	4.800	2.023	26	18	17.057	4.957
Plastic	3.380	1.425	16	20	11.468	4.090
Minerals	3.221	1.428	20	17	11.000	4.858
Basic metals	3.250	1.162	16	16	11.248	4.200
Misc. manuf.	3.926	1.023	18	17	12.265	3.059

<sup>a</sup>Millions of euros.

age will lead to differences in profits across firms. Differences in revenue, holding capital and age fixed, will be reflected in differences in productivity in the empirical model. In the model, capital, age, and productivity can all affect the expected benefit from investing in R&D. The second and third columns show that firms that invest in R&D generally have much larger capital stocks than firms that do not invest. The fourth and fifth columns indicate no substantial differences in median firm age between investing and noninvesting firms. In most industries, the firms that invest are equal in age or slightly older with the chemical, textile, and plastic industries being the exception. The last two columns of the table show that firms that invest in R&D have larger sales, suggesting there may be productivity differences between the two groups. Overall, the patterns in Table 2 suggest that firm capital and productivity are likely to be important sources of variation in the benefits of R&D across firms.

## 5. Empirical model

□ **Productivity evolution.** In this subsection, we describe how we use the MIP data to estimate the revenue function, the R&D-innovation, and the innovation-productivity relationships. Given that the innovation and R&D variables are discrete and observed in the data, we estimate the probability distribution  $F(d_{it+1}, z_{it+1} | rd_{it})$  as the fraction of observations reporting each of the four combinations of  $d_{it+1}$  and  $z_{it+1}$ , conditioning on  $rd_{it} = 0$  and  $rd_{it} = 1$ . The innovation probabilities are estimated separately for each of the 12 industries.

The demand elasticity for each industry is estimated using the fact that the model implies the ratio of total variable cost to firm revenue equals  $1 + 1/\eta$ . We use the mean variable cost-revenue ratio of each of the 12 industries as an estimate of one plus the inverse industry demand elasticity.

Estimates of the transition probabilities for productivity  $G(\omega_{it+1} | \omega_{it}, d_{it+1}, z_{it+1})$  are needed to construct the value function. Unlike the innovation and R&D variables, the firm's productivity is not observable, and the process of productivity evolution is estimated jointly with the parameters of the firm's revenue function,  $\beta_k$  and  $\beta_a$ , using the data on firm sales. To estimate the process of productivity evolution, we use the methodology developed by Doraszelski and Jaumandreu (2013). Using the structure of our model, we can solve for the demand functions for the variable inputs of labor and materials. The factor demand equation for the log of materials is:

$$m_{it} = \beta_l + (1 + \eta)\beta_k k_{it} + (1 + \eta)\beta_a a_{it} - (1 + \eta)\omega_{it}. \quad (10)$$

In this equation, the intercept  $\beta_t$  depends on the common time-varying components in the model, which include the intercept of the demand function and the variable input prices. The material demand depends on the observed capital stock, age, and unobserved firm productivity. Following the insight originally developed by Olley and Pakes (1996), this factor demand equation can be solved for the unobserved productivity as a function of the firm's material and capital inputs and age. Solving equation (10) for productivity and lagging it one period gives:

$$\omega_{it-1} = \left( \frac{1}{1+\eta} \right) \beta_{t-1} + \beta_k k_{it-1} + \beta_a a_{it-1} - \left( \frac{1}{1+\eta} \right) m_{it-1}. \quad (11)$$

We parameterize the productivity evolution process as a cubic function of lagged productivity and a full set of interactions between the dummy variables for product and process innovations:

$$\omega_{it} = \alpha_0 + \alpha_1 \omega_{it-1} + \alpha_2 \omega_{it-1}^2 + \alpha_3 \omega_{it-1}^3 + \alpha_4 d_{it} + \alpha_5 z_{it} + \alpha_6 d_{it} z_{it} + \varepsilon_{it}. \quad (12)$$

The persistence in firm productivity over time is captured by the coefficients  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$ . The effect of innovations on the mean of the distribution of future firm productivity is captured by the coefficients  $\alpha_4$ ,  $\alpha_5$ , and  $\alpha_6$ . The coefficient  $\alpha_6$  allows for the possibility that the marginal effect of either a product or process innovation on future productivity depends on whether the firm has the other type of innovation.<sup>18</sup> Substituting equation (11) into the productivity evolution equation (12) and that into the revenue function gives the estimating equation for firm revenue:

$$\begin{aligned} r_{it} = & \lambda_0 + \lambda_t + (1+\eta)\beta_k k_{it} - \alpha_1 [\beta_{t-1} + (1+\eta)\beta_k k_{it-1} + (1+\eta)\beta_a a_{it-1} - m_{it-1}] \quad (13) \\ & - \left( \frac{\alpha_2}{1+\eta} \right) [\beta_{t-1} + (1+\eta)\beta_k k_{it-1} + (1+\eta)\beta_a a_{it-1} - m_{it-1}]^2 \\ & - \left( \frac{\alpha_3}{(1+\eta)^2} \right) [\beta_{t-1} + (1+\eta)\beta_k k_{it-1} + (1+\eta)\beta_a a_{it-1} - m_{it-1}]^3 \\ & - (1+\eta)[\alpha_4 d_{it} + \alpha_5 z_{it} + \alpha_6 z_{it} d_{it}] - (1+\eta)\varepsilon_{it} + v_{it}. \end{aligned}$$

The error term  $v_{it}$  is a transitory shock to the firm's revenue function, which is not observed by the firm prior to choosing its variable inputs or making its R&D decision. For estimation, we utilize the moment conditions implied by the fact that the error term  $-(1+\eta)\varepsilon_{it} + v_{it}$  is uncorrelated with all right-hand side variables,  $a_{it-1}$ ,  $k_{it}$ ,  $k_{it-1}$ ,  $m_{it-1}$ ,  $z_{it}$ ,  $d_{it}$ , and  $z_{it} d_{it}$ . The intercept  $\lambda_0$  is a combination of the intercepts of the revenue function and the productivity evolution equation  $\alpha_0$ . We can separately identify the  $\alpha_0$  parameter from the revenue function intercepts using the moment condition that  $\varepsilon_{it}$  has a zero mean. The time coefficients  $\lambda_t$  and  $\beta_{t-1}$  are functions of the common time-varying variables, including the demand intercept and factor prices. The  $\beta_{t-1}$  coefficients are identified, up to a base-year normalization, and can be distinguished from the  $\lambda_t$  coefficients because of the higher-order powers on  $\omega_{it-1}$  in equation (12). We estimate equation (13) separately for the high-tech and low-tech industry groups using nonlinear least squares (NLLS). We allow the intercept  $\lambda_0$  to vary across the two-digit industries in each group, reflecting industry differences in the revenue functions, and include the industry-specific estimate of the demand elasticity as data. After estimation of the revenue function parameters, firm-level productivity is constructed from the inverted material demand function equation (11). The final parameter estimated is the variance of the error term in the productivity evolution equation  $\sigma_\varepsilon^2$ , and this is estimated from the residuals in the productivity evolution equation.

This process differs slightly from the methodology developed by Olley and Pakes (1996) in two respects. First, as in Doraszelski and Jaumandreu (2013) and Aw, Roberts, and Xu (2011), productivity evolution is not an exogenous process but is affected by the firm's R&D choice. In

<sup>18</sup> This interaction term also allows for the possibility that new product introductions may also require new production processes, which could potentially raise cost and reduce output, and thus offset some of the revenue gains from the product innovations. In this case, we may observe smaller productivity gains for firms reporting both innovations, relative to firms that report only one type of innovation.

this case, it is affected by the firm's innovations and, as a result, the innovation variables enter into equation (13). Second, because we are modelling productivity using the revenue function, we do not need to estimate the production function coefficients on the variable inputs of labor and materials. This simplifies equation (13) by removing the need to instrument variable input levels, which would appear on the right-hand side when using the production function as the starting point.

□ **Value function and the dynamic choice of R&D.** As described in Section 3, the firm bases its R&D investment decision on a comparison of the long-run payoff from undertaking R&D,  $\Delta EV(\omega_{it})$ , with the realized maintenance or startup cost,  $C_{it}$ . The probability that the firm chooses to invest in R&D is given by:

$$\begin{aligned} Pr(rd_{it} = 1|s_{it}) &= Pr[\Delta EV(\omega_{it}) \geq C_{it}(rd_{it-1})] \\ &= 1 - \exp(-\Delta EV(\omega_{it})/(\gamma^m * rd_{it-1} * k_{it} + \gamma^s * (1 - rd_{it-1}) * k_{it})), \end{aligned} \quad (14)$$

where the innovation cost is modelled as described in equation (6). In the empirical model, we allow the cost function parameters  $\gamma = (\gamma^m, \gamma^s)$  to vary by industry.

The final piece of the empirical model is the construction of the value function and  $\Delta EV(\omega_{it})$ , equations (7) and (9), respectively. We apply the nested fixed-point algorithm developed by Rust (1987) to estimate the dynamic discrete choice model. The state space  $s_{it} = (\omega_{it}, rd_{it-1})$  is discretized into 100 grid points for productivity and two values for lagged R&D choice, and we use value function iteration to solve for the value function at each element of this discretized state space. In addition, the firm value differs across firms based on capital stock, age, and industry. We define a set of 4800 discrete firm types based on 100 values of the capital stock, four age categories, and 12 industries, and construct the firm value and payoff to R&D for each firm type. The payoff to R&D is computed for each data point by using a cubic spline to interpolate across the productivity and capital grid points for each industry-age category.

Assuming the firm's state variables  $s_{it}$  are independent of the cost draws and that the costs are *i.i.d.* across all periods and firms, the likelihood function for the firms' R&D choice data is:

$$L(\gamma|rd, s) = \prod_i^N \prod_t^{T_i} Pr(rd_{it}|s_{it}; \gamma). \quad (15)$$

The vectors  $rd$  and  $s$  contain every firm's R&D choice and state variables for each period, respectively. The total number of firms is denoted by  $N$ , and  $T_i$  is the number of observations for firm  $i$ .

## 6. Empirical results

□ **Estimates of the innovation and productivity process.** Estimates of the probability of an innovation conditional on the firm's prior-period investment in R&D,  $Pr(d_{it+1}, z_{it+1}|rd_{it})$ , are reported for each industry in Table 3. There is a strong relationship between R&D investment and innovation outcomes. Columns (2) through (5) show the probability of realizing each combination of product and process innovation, given that the firm does not engage in R&D. Columns (6) through (9) report these probabilities for firms that conduct R&D. Column (2) shows that, on average, for firms that did not engage in R&D, the probability of having neither product nor process innovation in the next year averages 0.771 in the high-tech industries and 0.792 in the low-tech industries. This estimate is similar across all 12 industries, varying from a low of 0.710 in electronics to 0.820 in basic metals. What is more important to note is that approximately 22% of the firms still realize innovations even if their R&D spending is zero, and the most common outcome among the three combinations is the joint realization of product and process innovations ( $d = 1, z = 1$ ). This result indicates that prior-period R&D is not necessary for the firm to realize innovations.

**TABLE 3** Probability of Innovation Conditional on Past R&D:  $\Pr(d_{t+1}, z_{t+1} | rd_t)$ 

	$rd_t = 0$				$rd_t = 1$			
	$d = 0$	$d = 1$	$d = 0$	$d = 1$	$d = 0$	$d = 1$	$d = 0$	$d = 1$
Product Innovation	$z = 0$	$z = 0$	$z = 1$	$z = 1$	$z = 0$	$z = 0$	$z = 1$	$z = 1$
Process Innovation								
High-Tech Industries								
Chemicals	0.776	0.049	0.049	0.126	0.107	0.224	0.048	0.621
Machinery	0.780	0.061	0.036	0.123	0.104	0.252	0.038	0.606
Electronics	0.710	0.084	0.028	0.178	0.094	0.268	0.031	0.607
Instruments	0.800	0.050	0.020	0.130	0.091	0.302	0.011	0.595
Vehicles	0.778	0.065	0.040	0.111	0.135	0.174	0.058	0.633
<b>Average HT Industries</b>	<b>0.771</b>	<b>0.061</b>	<b>0.037</b>	<b>0.131</b>	<b>0.103</b>	<b>0.253</b>	<b>0.035</b>	<b>0.609</b>
Low-Tech Industries								
Food	0.756	0.047	0.047	0.150	0.239	0.178	0.046	0.537
Textiles	0.783	0.062	0.037	0.108	0.254	0.244	0.048	0.454
Paper	0.785	0.032	0.085	0.098	0.270	0.138	0.146	0.446
Plastic	0.793	0.077	0.017	0.111	0.145	0.171	0.040	0.644
Minerals	0.780	0.062	0.023	0.136	0.179	0.163	0.048	0.611
Basic metals	0.820	0.024	0.046	0.110	0.170	0.123	0.118	0.585
Misc. manuf.	0.780	0.083	0.038	0.098	0.167	0.259	0.051	0.523
<b>Average LT Industries</b>	<b>0.792</b>	<b>0.047</b>	<b>0.046</b>	<b>0.114</b>	<b>0.196</b>	<b>0.168</b>	<b>0.082</b>	<b>0.555</b>

Firms that invest in R&D are much more likely to report innovations. However, R&D is not sufficient to produce innovations: as shown in column (6), the probability of no innovation outcomes varies between 9.1% and 27.0% across the industries, and averages 0.196 in the low-tech industry group and 0.103 in the high-tech industries. The difference in the probability of innovation outcomes can reflect a combination of lower R&D effort in the low-tech industries, even when the firm reports positive R&D spending, and fewer technological opportunities for innovations. Among the three possible combinations of innovation outcomes, the most common is that the firm reports a product and process innovation ( $d = 1, z = 1$ ), with the probability varying between 0.446 and 0.644. On average, the probability of realizing both innovations is higher in the high-tech group than in the low-tech group, 0.609 versus 0.555. Among investing firms, the success rate for introducing a new product is in general higher than the rate for a new process. The only exception is the paper industry, which is an industry where large-scale production is important and which could give a strong incentive for firms to invest in process innovations.<sup>19</sup>

Table 4 reports the demand elasticity estimates for each industry in the high-tech and low-tech sectors. In the chemical industry, the estimate of  $(1 + 1/\hat{\eta})$  is 0.675, implying a demand elasticity  $\hat{\eta}$  of  $-3.075$ . The demand elasticity is used to convert productivity into profit, as seen in equations (3) and (4). The estimates vary substantially across industries, ranging from  $-2.991$  in the food industry to  $-5.266$  in metals.

Table 5 reports the estimates of the productivity evolution process for the high-tech and low-tech sectors from equation (13). The double and single asterisks denote parameter estimates different from zero at the 0.01 and 0.05 significance levels, respectively. The positive coefficient estimates for  $z$  and  $d$  indicate that firms that realize innovations have, on average, higher future productivity levels compared to those that do not have any kind of innovation. For firms in the high-tech industries, a new product innovation  $d$  contributes, on average, a 3.6% productivity gain, and a new process innovation  $z$  contributes 2.9%. There is no significant additional effect

<sup>19</sup> If we construct Table 3 using  $rd_{t-1}$  as the conditioning variable, so there is a two-year lag between R&D and innovation, we get a very similar pattern of innovation rates. Among the firms with  $rd_{t-1} = 0$ , 73.6% in high-tech and 76.4% in low-tech report no innovation. Among the firms with  $rd_{t-1} = 1$ , 22.5% in the low-tech sectors report no innovation, which is twice as large as in the high-tech sectors. The estimates of innovation probabilities by industry are not sensitive to the use of one- or two-period lags in R&D.

**TABLE 4 Demand Elasticity Estimates (standard error)**

	$1+1/\eta$	$\eta$	Sample Size
<b>High-Tech Industries</b>			
Chemicals	0.675 (0.005)**	-3.075	1350
Machinery	0.803 (0.002)**	-5.078	2643
Electronics	0.731 (0.005)**	-3.713	1403
Instruments	0.763 (0.005)**	-4.213	1428
Vehicles	0.796 (0.005)**	-4.891	892
<b>Low-Tech Industries</b>			
Food	0.665 (0.007)**	-2.991	1162
Textiles	0.697 (0.003)**	-3.302	991
Paper	0.697 (0.003)**	-3.296	1669
Plastic	0.798 (0.003)**	-4.941	1396
Minerals	0.675 (0.005)**	-3.080	960
Metals	0.810 (0.002)**	-5.266	2763
Misc. manuf.	0.765 (0.004)**	-4.253	872

\*\*Significant at the 0.01 level.

**TABLE 5 Productivity Evolution Parameters (standard error)**

	High-Tech Industries	Low-Tech Industries
$d$	0.036 (0.008)*	0.015 (0.008)
$z$	0.029 (0.014)*	0.035 (0.010)*
$d * z$	0.001 (0.016)	-0.009 (0.013)
$\omega_{t-1}$	0.711 (0.020)*	0.707 (0.015)*
$\omega_{t-1}^2$	0.211 (0.012)*	0.160 (0.010)*
$\omega_{t-1}^3$	-0.056 (0.004)*	-0.041 (0.003)*
$k$	-0.065 (0.003)*	-0.058 (0.003)*
$a = (10 - 19)$	0.009 (0.013)	0.017 (0.013)
$a = (20 - 49)$	-0.058 (0.019)*	-0.049 (0.020)*
$a \geq 50$	-0.158 (0.025)*	-0.101 (0.023)*
<i>Intercept</i>	0.866 (0.150)*	0.313 (0.182)*
<i>Chemicals</i>	0.061 (0.037)	0.022 (0.037)
<i>Machinery</i>	0.035 (0.030)	-0.162 (0.036)*
<i>Electronics</i>	0.069 (0.034)*	-0.033 (0.033)
<i>Instruments</i>	0.072 (0.033)*	-0.053 (0.033)
		<i>Minerals</i>
		-0.001 (0.037)
		<i>Metals</i>
		0.006 (0.031)
$SE(\epsilon)$	0.189	0.203
Sample size	3313	4290

\*Significant at the 0.05 level, \*\*Significant at the 0.01 level.  
Time dummies for  $\lambda_t$  and  $\beta_{t-1}$  included but not reported.

from having both types of innovations jointly. The coefficient on the interaction term  $d * z$  is 0.001, which implies an average productivity increase of 6.6% in the next year if the firm has both types of innovations.

In the low-tech industries, the magnitudes of these productivity effects are smaller, but their difference is more pronounced. Firms that introduced a new product have, on average, 1.5% higher future productivity, whereas a new process innovation raises productivity by 3.5%. One possible reason for the weaker impact of product innovation on future productivity is that new or improved products may represent less substantial changes over existing products in these industries.<sup>20</sup> If a firm realizes product and process innovations, it has 4.1% higher future productivity.

<sup>20</sup> This interpretation is supported by data on products that are new to the market. In the MIP, the proportion of firms introducing products that are new to the market varies from 39% to 51% in the high-tech industries compared to 16% to 31% in the low-tech industries.



The effect of past productivity on the current productivity level is measured by the coefficients of  $\omega_{t-1}$ , its squared and cubic terms. Past productivity is highly persistent. There is a nonlinear relationship between current and lagged productivity for both sectors, as seen by the statistically significant effect of  $\omega_{t-1}^2$  and  $\omega_{t-1}^3$ . The persistence of the productivity process has a substantial impact on the long-run payoff from R&D because it determines how quickly the productivity gains from an innovation depreciate. Lower values of  $\alpha_1$  imply more rapid depreciation of the productivity and profit gains from an innovation  $d$  or  $z$ , which lowers the long-run payoff to R&D. Overall, large coefficients on the innovation variables and high persistence of the productivity process results in a high level of  $\Delta EV$ , the expected long-run payoff to R&D.

The remaining variables in the profit function, capital and age, also have significant effects. The cost elasticity of capital in the high-tech sector is estimated to be  $\hat{\beta}_k = -0.065$  and  $-0.058$  in the low-tech sector. Negative values of  $\beta_k$  imply firms with a higher capital stock have lower production costs because they use fewer variable inputs. We include three dummy variables representing different age categories in the estimation, where firms between one and nine years old are the omitted base group. The estimated coefficients for firms in the 10–19 category implies that they have higher costs than the base group, but the difference is not significant in either industry. The coefficients on the remaining two age groups are negative and statistically significant, indicating that older firms have lower costs than the base group. The magnitude of the estimates is higher for the oldest firms, indicating higher cost efficiency as the firm ages.<sup>21</sup>

Before turning to the estimation of the dynamic parameters, we examine the reduced-form relationship between R&D, the firm characteristics, capital, age, and industry, and the two state variables, productivity and lagged R&D. Table 6 reports probit regressions of this relationship. The second column of the table reports estimates for the high-tech industries without the lagged R&D variable included. Capital and productivity are significant determinants of discrete R&D choice, but age is not significant. When lagged R&D is included in the specification, it is, not surprisingly, highly significant and the other coefficients drop substantially in magnitude, but capital and productivity remain statistically significant. A similar pattern is found for the low-tech industries. The main difference is that productivity is not significant when lagged R&D is included. Also, the age coefficients are more important than in the high-tech industries and the negative signs imply that older firms are less likely to invest in R&D. This age pattern is not consistent with the role of age in the dynamic model, where we find that older firms have lower production costs which will tend to increase their incentive to invest in R&D.

□ **Estimates of the cost of innovation.** The final sets of parameter estimates characterize the innovation cost distributions. In the model, this is the expenditure the firm must incur to generate a product or process innovation that, in turn, raises its productivity. Table 7 reports parameter estimates for three different cost specifications. The standard errors are bootstrapped to account for the variation in the first-stage parameter estimates. The first, labelled model A in the table, is the specification described in equation (6), which allows the mean of the cost distribution to differ with the firm's capital stock and industry. Model B allows the mean of the cost distribution to differ by industry, whereas model C allows it to differ across small-, medium-, and large-size

<sup>21</sup> We also check if estimates of the revenue function and productivity process are affected by firm exit. Olley and Pakes (1996) found that the use of a balanced panel of plants did result in selection bias in their production function estimates, but once the observations on entering and exiting plants were included in the sample, the selection bias was very minor. As explained in the first subsection of Section 4, virtually all of the attrition in our sample is due to nonreporting, not firm death. To verify if sample attrition affects our results, we estimate probit models of exit and find that capital stock, age, material expenditures, and industry have no explanatory power in the regressions. The pseudo  $R^2$  in the exit regression is 0.005 in the high-tech and low-tech models, with only a few industry dummies and age dummies being close to statistically significant. Including the predicted probability of exit in the revenue function estimation, equation (13), has no effect on the structural parameter estimates. The estimates of  $\omega$  from a model that controls for the probability of exit and a model that does not have a correlation above .99. There is no evidence of selection bias from sample attrition in our estimates.

TABLE 6 Reduced-Form Model of R&amp;D Investment

	High-Tech Industries		Low-Tech Industries	
$rd_{t-1}$		2.048 (0.065)*		1.608 (0.046)*
$k$	0.192 (0.017)*	0.118 (0.020)*	0.167 (0.014)*	0.088 (0.016)*
$k^2$	0.003 (0.005)	0.002 (0.006)	0.014 (0.004)*	0.011 (0.005)*
$\omega$	0.313 (0.066)*	0.163 (0.081)*	0.130 (0.054)*	0.013 (0.060)
$\omega^2$	-0.203 (0.039)*	-0.134 (0.047)*	-0.062 (0.028)*	0.004 (0.032)
$a = (10 - 19)$	-0.003 (0.068)	0.023 (0.083)	-0.091 (0.054)	-0.101 (0.061)
$a = (20 - 49)$	-0.083 (0.071)	-0.099 (0.086)	-0.232 (0.059)*	-0.141 (0.066)*
$a \geq 50$	0.041 (0.080)	-0.028 (0.096)	-0.155 (0.056)*	-0.093 (0.063)
<i>Intercept</i>	0.457 (0.084)*	-0.729 (0.110)*	<i>Intercept</i>	-0.062 (0.075)
<i>Chemicals</i>	0.090 (0.081)	0.026 (0.101)	<i>Food</i>	0.020 (0.087)
<i>Machinery</i>	0.461 (0.092)*	0.254 (0.112)*	<i>Textiles</i>	-0.110 (0.078)
<i>Electronics</i>	0.568 (0.172)*	0.248 (0.115)*	<i>Paper</i>	0.336 (0.085)*
<i>Instruments</i>	-0.195 (0.100)	-0.112 (0.123)	<i>plastic</i>	0.149 (0.088)
			<i>Minerals</i>	0.197 (0.077)*
			<i>Metals</i>	0.346 (0.096)*
Sample size	3313	3313	4290	4290
Pseudo $R^2$	0.104	0.422	0.070	0.301

\*Significant at the 0.05 level, \*\*Significant at the 0.01 level.

classes defined by the firm's capital stock. All three cost specifications distinguish between the distribution of startup and maintenance costs.

Focusing on specification A, the estimated maintenance costs are smaller than startup costs in all industries. This means that, comparing two firms with the same productivity, capital stock, age, and industry, and, hence the same expected payoff to R&D, the firm with previous R&D experience will find it less expensive to generate an innovation than one with no prior experience.

In specification A, the positive coefficient implies that firms with larger capital stock draw their innovation cost from a distribution with a higher mean. In the estimated model, the payoff to conducting R&D increases with the capital stock. Holding productivity, age, and industry fixed, a firm with a larger capital stock will have a larger benefit to investing in R&D and will be willing to invest more resources to get an innovation. An analogous interpretation of the cost magnitude applies to a comparison of the high-tech and low-tech industries. The higher expected payoff to R&D in the high-tech industries allows these firms to invest more in the innovation process.

The two other cost specifications result in lower log-likelihood values. The results for model B show that there are significant differences in mean costs across industries. These differences are particularly large between the high-tech and low-tech industries. The estimates for model C show a higher mean level of innovation costs for larger firms. Both of these patterns, industry differences and variation with firm size, are incorporated in model A and we will focus on those results in the remainder of this article.

We assess the goodness of fit of the dynamic model by simulating the firms' investment choices given their capital stock, age, productivity, and industry and compute the predicted R&D transition rates. Table 8 reports the actual and predicted R&D transition rates, using the cost estimates from model A. The model fits the data well for the firms that do not engage in R&D. In the high-tech industry, the predicted startup rate is 0.291, whereas this rate is 0.245 in the data. In the low-tech industry, the predicted rate is 0.233 and the observed rate is 0.219. For investing firms in the high-tech industry, our model predicts a continuation rate of 0.862 versus 0.933 in the data. For the low-tech industries, the prediction is 0.711 and the data is 0.824. Overall, the model slightly underestimates the continuation rate and overestimates the exit rate.

□ **Expected benefits of R&D.** Using the estimates of innovation costs, innovation probabilities, and the productivity process, we construct  $\Delta EV(\omega)$ , the expected long-run payoff to investing

**TABLE 7 Dynamic Cost Estimates (bootstrap standard errors)**

	Cost Model A		Cost Model B		Cost Model C	
	Startup	Maintenance	Startup	Maintenance	Startup	Maintenance
High-Tech Industry						
Chemicals	1.445 (0.5791)	0.254 (0.0872)	14.222 (6.6522)	2.031 (0.8472)	3.960 (2.1744)	0.526 (0.2700)
Machinery	1.494 (0.7766)	0.238 (0.1347)	9.302 (4.6963)	1.696 (0.9407)	11.421 (4.2505)	1.568 (0.6519)
Electronics	1.286 (0.8111)	0.097 (0.0518)	8.746 (5.9190)	0.646 (0.3211)	21.934 (7.9417)	3.730 (1.5551)
Instruments	0.758 (0.4273)	0.096 (0.0458)	3.536 (1.9357)	0.526 (0.2539)		
Vehicles	1.704 (1.0928)	0.405 (0.2024)	13.536 (8.3145)	3.533 (1.7479)		
Log-likelihood function (LLF)		-1728.6		-1935.6		-1793.7
Low-Tech Industry						
Food	0.317 (0.1490)	0.097 (0.0312)	2.722 (1.3074)	0.865 (0.2852)	1.538 (0.5299)	0.351 (0.1517)
Textiles	0.449 (0.2440)	0.084 (0.0290)	3.511 (1.7861)	0.733 (0.2720)	4.114 (1.4585)	0.827 (0.2930)
Paper	0.377 (0.1550)	0.091 (0.0338)	3.254 (1.3591)	0.810 (0.2988)	9.090 (3.1122)	1.596 (0.4986)
Plastic	0.640 (0.2660)	0.117 (0.0429)	4.305 (1.6991)	0.889 (0.3260)		
Minerals	0.307 (0.1446)	0.058 (0.0212)	2.324 (0.9807)	0.464 (0.1771)		
Metals	1.098 (0.5517)	0.162 (0.0743)	7.861 (3.8119)	1.277 (0.5738)		
Misc. manuf.	0.386 (0.2490)	0.067 (0.0281)	2.738 (2.3833)	0.503 (0.2181)		
LLF		-2913.1		-3194.2		-2946.5

LLF: log-likelihood function.

TABLE 8 Transition Rates for R&amp;D

	Data		Model	
	$rd_t = 0$	$rd_t = 1$	$rd_t = 0$	$rd_t = 1$
High-Tech Industries				
$rd_{t-1} = 0$	0.755	0.245	0.709	0.291
$rd_{t-1} = 1$	0.067	0.933	0.138	0.862
Low-Tech Industries				
$rd_{t-1} = 0$	0.781	0.219	0.767	0.233
$rd_{t-1} = 1$	0.175	0.824	0.289	0.711

TABLE 9 Long-Run Expected Benefits of Conducting R&D (evaluated at median  $\omega$ , median  $k$ , age = 2 in each industry)

	$\Delta EV(\omega)^a$	$Pr(rd_t = 1)$	
		$rd_{t-1} = 1$	$rd_{t-1} = 0$
High-Tech Industries			
Chemicals	4.213	0.847	0.281
Machinery	4.708	0.917	0.327
Electronics	2.691	0.975	0.244
Instruments	2.331	0.965	0.345
Vehicles	6.770	0.856	0.370
<b>Average HT Industries</b>	<b>4.413</b>	<b>0.912</b>	<b>0.313</b>
Low-Tech Industries			
Food	0.470	0.440	0.163
Textiles	0.299	0.397	0.090
Paper	0.450	0.459	0.138
Plastic	1.272	0.755	0.226
Minerals	0.590	0.730	0.218
Metals	1.563	0.716	0.170
Misc. manuf.	0.833	0.804	0.246
<b>Average LT Industries</b>	<b>0.782</b>	<b>0.614</b>	<b>0.179</b>

<sup>a</sup>Millions of euros.

in R&D, from equation (9). This measures the difference in the present value of expected future profits that accrue to the firm if it engages in R&D versus if it does not engage in R&D for a given year. This benefit depends on the industry-level measures (profit function, demand elasticity, productivity evolution, and innovation probabilities) and the firm-level variables (productivity, capital stock, and age) and, therefore, varies across firms within an industry.

Table 9 focuses on differences in the expected benefits across industries. The second column reports the value of  $\Delta EV(\omega)$  for a firm in age group 2 (10–19 years old) with the median level of productivity and capital stock in each industry, a “median” firm. In the high-tech industries, the expected payoff to R&D varies from 2.33 million euros in the instrument industry to 6.77 million euros in vehicles. Not surprisingly, given the earlier findings of lower innovation rates and lower productivity impacts of innovation, the expected payoff to R&D investment by firms in the low-tech sector is smaller. They vary from 0.299 million euros in the textile industry to 1.563 million in the basic metals industry.

These differences in expected benefits, when combined with the cost estimates, translate into differences in the probability of investing in R&D. The last two columns of Table 9 summarize these probabilities for the same “median” firm in each industry. The probabilities also vary with the firm’s prior R&D experience because of the difference in maintenance versus startup cost. In the high-tech industry, this “median” firm has a high probability of continuing to invest in R&D. Given they are paying a maintenance cost (i.e.,  $rd_{t-1} = 1$ ), the probability of investing in

R&D varies between 0.847 and 0.975. The probability of beginning an R&D program is much smaller, varying from 0.244 to 0.370 across industries, reflecting the substantial startup costs that are faced by firms beginning to invest in R&D. In the low-tech industries, continuing and starting probabilities are smaller, reflecting the lower expected payoffs. For the “median” firm with prior R&D experience, the investment probability varies from 0.397 to 0.804 and is particularly low for the food, textile, and paper industries. The probabilities of starting R&D vary from 0.09 to 0.246 across industries. This indicates a substantial hurdle to R&D startup in these industries.

Table 10 focuses on how the other firm characteristics, productivity, capital stock, and age, lead to variation in the expected benefit of R&D within each industry. The marginal effects of changes in productivity, capital, or age on  $\Delta EV(\omega)$  and the probabilities of investing in R&D are difficult to gauge directly from the structural coefficients in Tables 3, 4, 5, and 7 because they are nonlinear functions of model parameters and variables. In this table, we report estimates of the *change* in  $\Delta EV(\omega)$  and investment probabilities resulting from changes in the firm’s productivity, capital, and age. Columns (2), (3), and (4) summarize the difference in outcomes between firms with productivity levels at the 25th and 75th percentile of the productivity distribution, holding capital stock fixed at the median level and age at category 2. Productivity impacts the firm’s expected return and probability of investment directly through the profit function but also through the persistence parameters in the productivity evolution process. High persistence, which we observe in the Table 5 estimates, implies long-lived productivity differences across firms. Therefore, a high-productivity firm is more likely to invest in R&D than a low-productivity firm.

The results in columns (2)–(4) show that productivity heterogeneity is a major driving force of the difference in expected returns and investment probability across firms. In the high-tech industries, a firm at the 75th percentile of the productivity distribution has an expected payoff that is 3.5 to 15.3 million euros higher than that of a firm at the 25th percentile. The average difference across industries is 8,899 million euros. There are also substantial differences in the probability of R&D investment. In the chemical industry, the probability of investing is 0.686 (0.565) higher for a high-productivity firm paying a maintenance (startup) cost compared to a low-productivity firm. The increase in investment probability is also substantial in the vehicle industry. In the low-tech industries, the magnitude of the differences in expected returns is much smaller, varying from approximately 1.0 to 2.5 million euros and averaging 1.608 million. The smaller magnitude reflects lower overall expected returns in these industries. Despite a small difference in expected returns, we observe a relatively large impact of productivity changes on the probability of investing, particularly on the probability of continuing to invest. This probability increase averages 0.594 across the seven industries, whereas the startup probability is higher by an average of 0.329.

Columns (5), (6), and (7) provide the estimated impacts of changes in the firm’s capital stock on expected payoffs and investment probabilities. These reported numbers summarize the difference in outcomes between firms at the 25th and 75th percentile of the capital stock distribution, holding productivity fixed at the median level and age fixed at group 2. Capital differences impact these outcomes through two channels: through the firm’s short-run profit function and through the cost of innovation. Although capital has a positive impact on profits, it also increases innovation costs, and these effects are offsetting on the probability of investing. When compared with productivity heterogeneity, the differences across firms’ capital stocks account for much smaller differences in expected returns and investment probabilities. This is true in the high-tech and low-tech industries. Among the high-tech industries, the increase in expected return averages 2.943 million euros and the probabilities of investing in R&D increase by 0.062 and 0.080, on average, for the maintenance and startup groups, respectively. In the low-tech industries, the average difference in expected returns is 0.443 million euros and the probabilities differ by 0.043 and 0.027. In particular, in four of the low-tech industries, the difference in investment probabilities between firms with large and small capital stocks is less than one percentage point and is even negative in some cases.

TABLE 10 Within-Industry Differences in the Benefits of R&D and Probability of Investing

	Impact of Increasing $\omega$ on		Impact of Increasing $k$ on		Impact of Increasing Age on	
	$Pr(rd_t = 1)$		$Pr(rd_t = 1)$		$Pr(rd_t = 1)$	
	$rd_{t-1} = 1$	$rd_{t-1} = 0$	$rd_{t-1} = 1$	$rd_{t-1} = 0$	$rd_{t-1} = 1$	$rd_{t-1} = 0$
<b>High-Tech Industries</b>	$\Delta EV(\omega)^*$		$\Delta EV(\omega)^*$		$\Delta EV(\omega)^*$	
Chemicals	11.815	0.565	1.594	0.017	2.415	0.124
Machinery	6.596	0.342	3.753	0.096	5.345	0.244
Electronics	7.240	0.185	1.404	0.013	1.960	0.139
Instruments	3.563	0.154	1.450	0.026	1.758	0.179
Vehicles	15.282	0.542	6.512	0.154	7.665	0.257
Average HT industries	<b>8.899</b>	<b>0.364</b>	<b>2.943</b>	<b>0.062</b>	<b>3.829</b>	<b>0.188</b>
<b>Low-Tech Industries</b>						
Food	1.742	0.460	0.122	-0.007	0.146	0.045
Textiles	1.015	0.266	0.114	-0.012	0.116	0.033
Paper	1.264	0.698	0.165	0.004	0.173	0.048
Plastic	2.005	0.424	0.880	0.116	1.039	0.146
Minerals	1.257	0.649	0.179	-0.008	0.220	0.069
Metals	2.531	0.458	1.162	0.144	1.527	0.138
Misc. manuf.	1.444	0.476	0.475	0.063	0.523	0.122
Average LT industries	<b>1.608</b>	<b>0.594</b>	<b>0.443</b>	<b>0.043</b>	<b>0.535</b>	<b>0.086</b>

The final three columns examine the marginal effect of age. We compare the outcomes of a firm in age group 4 ( $\geq 50$  years) to the outcomes of a firm in age group 2 (10–19 years) holding capital and productivity fixed at the median industry values. Firm age has an impact on the expected returns and investment rates through its impact on the firm's profit. The differences in outcomes between age groups are slightly larger than those observed for capital heterogeneity but much less than those for productivity heterogeneity. Averaging across industries, the mean expected benefit of an older firm is 3.829 and 0.535 million euros higher in the high-tech and low-tech groups, respectively. In the high-tech industry, the probability of starting to invest differs more across the age categories than the probability of maintaining investment. In the low-tech industries, the opposite pattern is observed. Age differences have a larger impact on the probability of continuing to invest.

Overall, the results in Tables 9 and 10 show that there are substantial differences in the expected return to R&D, particularly between firms in the low-tech and high-tech industries. This contributes to the differences in R&D investment rates across industries. Within industry, heterogeneity in productivity is particularly important in explaining differences across firms, with high-productivity firms having substantially higher expected returns and investment probabilities. Differences in capital stocks and age also contribute to within-industry differences in these outcomes; however, the magnitude of their impact is smaller than the impact of productivity differences.

□ **The long-run and short-run return to R&D.** The dynamic framework developed in this article has the advantage of providing measures for the long-run and short-run benefits of R&D investment that can be compared to each other. The short-run gain captures changes in sales and profits in the subsequent period, whereas the long-run gain captures the changes in firm value due to the firm being on a higher productivity path. The latter includes a higher profit stream and different optimal future R&D choices. Both of these effects are induced by the productivity gain resulting from R&D investment.

We define the long-run gain as the proportional impact of R&D on firm value. It is measured as the log difference in the expected future value of the firm, equation (8), conditional on its R&D choice, while holding the firm's other characteristics fixed:

$$\Delta \ln EV = \ln(EV(s_{it+1}|\omega_{it}, rd_{it} = 1)) - \ln(EV(s_{it+1}|\omega_{it}, rd_{it} = 0)). \quad (16)$$

This long-run gain can be constructed for every firm-year observation in the data, regardless of whether or not the firm actually invested in R&D in that year. This allows us to characterize the distribution of expected long-run gains over all firms, and not simply for firms that choose to conduct R&D. The median value over all sample observations in each industry is reported in the second column of Table 11, the 25th and 75th percentiles are reported in column (3). This table also reports the median value of  $EV(s_{it+1}|\omega_{it}, rd_{it} = 0)$ , denoted  $EV0$ , which serves as the base for interpreting the proportional change in firm value.

In the high-tech industries, the median value of the long-run gain varies between 0.057 in the instrument industry to 0.085 in vehicles. The 25th percentile varies between 0.024 and 0.048, whereas the 75th percentile varies from 0.079 to 0.109. Aggregating over all firms in the high-tech industries, the median gain is 0.067, implying a difference in long-run firm value of 6.7% between firms that undertake R&D and those that do not. The 25th and 75th percentiles are 0.040 and 0.093, with much of the heterogeneity arising within industries.

In the low-tech industries, the proportional long-run gains are lower. The median varies from a low of 0.017 in the food and textile industries to 0.390 in the metals industry and equals 0.028 when aggregating over firms in the low-tech industries. In three of the industries, the median gain is less than 2.0% of firm value. The dispersion in long-run benefits within each industry is also much smaller than in the high-tech industries. In four of the industries, the 25th percentile is less than or equal to 0.01, and the 75th percentile is above 0.05 in only two cases, plastics and metals

TABLE 11 Long-Run and Short-Run Proportional Gains from R&amp;D

	Long Run, $\Delta \ln EV$			Short Run, $\Delta r$	
	Median	25th–75th	Median $EV0$	$\Delta r$	Median Revenue
<b>High-Tech Industries</b>					
Chemicals	0.067	0.024–0.091	71.87	0.081	31.44
Machinery	0.071	0.048–0.091	80.06	0.158	12.40
Electronics	0.068	0.039–0.098	43.84	0.095	10.23
Instruments	0.057	0.037–0.079	45.79	0.127	6.21
Vehicles	0.085	0.040–0.109	104.88	0.151	30.11
<b>All HT Industries</b>	<b>0.067</b>	<b>0.040–0.093</b>	<b>64.59</b>	<b>0.122</b>	<b>20.68</b>
<b>Low-Tech Industries</b>					
Food	0.017	0.007–0.030	30.13	0.035	12.78
Textiles	0.018	0.007–0.034	21.58	0.040	7.31
Paper	0.019	0.007–0.033	27.18	0.041	9.60
Plastic	0.037	0.023–0.056	37.67	0.095	6.90
Minerals	0.025	0.010–0.038	26.06	0.046	7.58
Metals	0.039	0.023–0.059	49.41	0.100	7.16
Misc. manuf.	0.032	0.016–0.045	30.90	0.067	8.47
<b>All LT Industries</b>	<b>0.028</b>	<b>0.013–0.041</b>	<b>34.25</b>	<b>0.061</b>	<b>7.82</b>

industries. Not surprisingly, given the results reported in Tables 3, 4, and 5, R&D investment has a lower expected payoff in the low-tech industries relative to the high-tech industries.

As discussed in Section 2, the knowledge production framework focuses on the elasticity of output (usually measured as firm revenue) with respect to R&D expenditure as a measure of the return to additional R&D spending. Hall, Mairesse, and Mohnen (2010) report that revenue elasticity estimates vary across studies from 0.01 to 0.25 and are centered around 0.08. Doraszelski and Jaumandreu (2013) report estimates of the elasticity of output, not revenue, for 10 Spanish manufacturing industries. The average value over all firms is 0.015, and the average at the industry level varies from  $-0.006$  to  $0.046$  across the 10 industries, with half of the industries falling between 0.013 and 0.022.

Using the results reported in Tables 3, 4, and 5, we construct an analogous measure using the discrete R&D variable: the proportional gain in firm revenue resulting when the firm moves from not investing in R&D ( $rd_t = 0$ ) to investing in R&D ( $rd_t = 1$ ). The revenue increase resulting from R&D depends on how R&D affects innovation, how innovation affects productivity, and how productivity translates into revenue. The difference in log revenue when  $rd = 1$  and  $rd = 0$  can be measured using our model as:

$$\Delta r = -(1 + \eta) \sum_{(d,z)} [g(\omega, d, z) - g(\omega, 0, 0)] [\Pr(d, z|rd = 1) - \Pr(d, z|rd = 0)] \quad (17)$$

for all  $(d, z) \in \{(1, 0), (0, 1), (1, 1)\}$ .

Table 11 provides estimates of this shift on the log of future revenue for each industry. These estimates are constant across firms in an industry. The final column of the table reports the median value of firm sales, in millions of euros, to use as the basis for comparison. In the high-tech industries,  $\Delta r$  varies from 0.081 in chemicals to 0.158 in machinery, and averages 0.122 across the five industries. The latter number implies 12.2% higher revenue for firms that conduct R&D relative to firms that do not. In the low-tech industries, the short-run revenue differences vary from 0.035 to 0.100 and average 0.061. These are within the range reported by Hall, Mohnen, and Mairesse (2010) in their review of the literature. In percentage terms, the short-run gains are larger in magnitude than the median long-run gains reported in column (2). However, the long-run gains  $\Delta \ln EV$  apply to a larger base, the expected future firm value, than the short-run gains in revenue. In monetary units, the short-run gains are always smaller than the long-run gains. This



TABLE 12 Effect of Technology Parameters on Productivity and R&amp;D

	HT industries with LT parameters			LT industries with HT parameters		
	Productivity Evolution	Innovation Function	Both	Productivity Evolution	Innovation Function	Both
Change in R&D Prop.						
5 years	-0.312	-0.080	-0.350	0.252	0.025	0.317
10 years	-0.375	-0.093	-0.412	0.283	0.021	0.362
Prop. Change in $\omega$						
5 years	-0.063	-0.019	-0.072	0.051	0.001	0.069
10 years	-0.104	-0.033	-0.111	0.082	0.009	0.110
Change in $\Delta EV$						
5 years	-5.490	-1.825	-6.035	2.069	0.033	2.481
10 years	-4.986	-1.765	-5.378	2.041	0.011	2.598

results because  $\Delta r$  does not consider the persistence of the productivity gains and the optimal future investment choices motivated by this productivity gain.

## 7. Counterfactual analysis

■ The results in the previous section reveal large differences in the long-run payoff to R&D, which generates differences in the investment incentive between firms in the low-tech and high-tech industries. We explore the source of these differences by focusing on the two technology components, productivity evolution and the innovation probability, that contribute to  $\Delta EV$ .<sup>22</sup> When investing in R&D, firms in the high-tech industries have a higher innovation success rate and their innovations have a larger productivity impact than those in the low-tech industries, as seen in Tables 3 and 5. Columns (2)–(4) of Table 12 report the average change in high-tech firm R&D investment, productivity level, and  $\Delta EV$  when facing low-tech productivity impact and innovation process. In column (2), we show how the high-tech firms would be affected if they had the  $\alpha_4$ ,  $\alpha_5$ , and  $\alpha_6$  parameters of the low-tech industries. The proportion of firms investing in R&D would drop by 0.312 after 5 years and 0.375 after 10 years. This reduction in R&D investment shows up very quickly in response to the reduced economic impact of innovations. The strong decline in R&D investment is followed by an increasing shortfall in the level of industry productivity. Average industry productivity drops by 6.3% after 5 years and this drop continues to fall to 10.4% after 10 years. The expected benefit of R&D decreases by approximately 5.0 million euros due to the lower productivity and smaller impact of innovations.

The third column of Table 12 shows the effect of lower innovation probabilities. If high-tech firms faced the same probabilities as in the low-tech sector, the R&D investment rate would drop by 0.080 and 0.093 over 5- and 10- year horizons, respectively. As a consequence, mean productivity level declines by 1.9% after 5 years and 3.3% after 10 years. Comparing the effects of these two technology factors on industry investment and productivity, the productivity impact of innovations is the stronger driving force for the difference between the high-tech and low-tech industry. Finally, the fourth column shows declines in R&D participation, mean productivity level, and expected benefits if productivity impact and innovation probabilities were identical to those of the low-tech sector. The declines are only slightly higher than those due to lower productivity impact of innovations, reflecting the nonlinear effect of these factors in our model.

<sup>22</sup> In the simulation exercises reported in this section, all observations are initialized with the observed industry, age, capital stock, productivity, and lagged R&D status in the data. The firm's cost, productivity shock, and innovation outcome are then drawn from the appropriate distributions and the firm's optimal R&D choice is calculated. The state variables are updated and the process is repeated for 10 periods. The simulations are repeated 100 times and averaged. All figures reported in Tables 12 and 13 are the differences in the simulated outcomes under alternative parameter choices relative to the base case, using the estimated model parameters.

TABLE 13 Counterfactual Reductions in Costs of Innovation

	20% Reduction in Maintenance Cost		20% Reduction in Startup Cost	
	High-Tech Ind.	Low-Tech Ind.	High-Tech Ind.	Low-Tech Ind.
Change in R&D Prop.				
5 years	0.079	0.070	0.002	0.013
10 years	0.090	0.072	0.004	0.012
Prop. Change in $\omega$				
5 years	0.008	0.005	0.002	0.021
10 years	0.014	0.006	0.003	0.070
Change in $\Delta EV$				
5 years	0.506	0.156	-0.455	-0.104
10 years	0.556	0.156	-0.373	-0.081

The last three columns of Table 12 report corresponding changes in R&D rate, productivity, and expected benefit for firms in the low-tech industries when imposing the high-tech innovation probabilities and productivity impact of innovations. Similar to the results found for high-tech industries, the major impact comes from the productivity evolution parameters. The stronger impact of innovation in the productivity process causes the R&D rate to increase by 0.252 after 5 years and 0.283 after 10 years. More favorable innovation probabilities, such as those faced by the high-tech firms, however, only increase the R&D rate by 0.025 and 0.021, respectively. Consequently, productivity changes attributed to the higher innovation impact are more pronounced than those resulting from higher innovation probabilities. In summary, Table 12 shows that the impact of innovation on firm productivity, and thus on firm sales and profits, plays a crucial role in explaining the differences in firm performance and R&D investment between high-tech and low-tech industries, more so than the impact of R&D on the innovation creation.

Another important component determining firm investment decisions is the cost of innovation. Policy instruments such as R&D subsidies, grants, and tax relief directly alter this cost. To assess the impact of tax relief or subsidies, we simulate firm R&D investment choices and productivity development if there was a permanent reduction in maintenance or startup costs. The second and third columns of Table 13 show the effect of lowering the maintenance cost distribution parameter by 20% for high-tech and low-tech industries. We report these effects as changes in R&D investment rates, mean productivity, and expected benefits  $\Delta EV$  after 5 and 10 years. The last two columns report the corresponding changes resulting from a 20% reduction in startup costs.<sup>23</sup>

Five years after a maintenance cost reduction, the probability of investing in R&D has risen, on average across the high-tech firms by 7.9 percentage points, and by 7.0 among the low-tech industry firms. After 10 years, we see a small additional increase in the proportion of firms investing, however, the largest increase in the investment rate resulted shortly after the cost reduction. In addition, the cost reduction does not affect all firms equally. Approximately 25% of the firms in each industry group are not affected by the maintenance cost reduction because their probability of investing was already close to either one or zero. The reduction in cost alters the investment decision of firms that have costs near the threshold of  $\Delta EV$ , whereas the decision of firms with very low or very high initial costs remains unchanged.

The increased rate of investment in R&D leads to a shift in the distribution of productivity. In the high-tech industries, the mean of the firm productivity distribution increases by 0.8% and 1.4% after 5 and 10 years, respectively. The increases are smaller, 0.5% and 0.6% in the

<sup>23</sup> In our model, there are no R&D spillovers across firms and the benefit of a firm's R&D investment is fully internalized by the firm, so there is no reason that a subsidy is needed to induce the socially efficient level of R&D investment. This counterfactual is simply summarizing how responsive firm R&D participation and productivity are to a reduction in innovation costs. See Klette, Moen, and Griliches (2000) for a review of the literature on subsidies and market failure.

low-tech industries. The lower cost of an innovation also raises the expected long-run benefit of investing in R&D. In the high-tech industries, the mean value of  $\Delta EV$  rises by slightly more than 0.5 million euros, whereas in the low-tech industries, it increases by 0.156 million euros.

The effects of a reduction in startup costs differ from what we observe for the maintenance cost reduction. A startup cost reduction lowers the entry barrier into investment for firms with no formal R&D previously. The results in Table 13 show there is virtually no effect of this cost reduction on R&D investment or productivity in the high-tech industries and a modest positive effect in the low-tech industries. In contrast to the maintenance cost reduction, which always raises the expected benefit of investment and increases the firm's probability of investing, a startup cost reduction has two opposing effects on R&D choice, which can explain the smaller impact. First, it lowers the entry cost for firms that were not investing, and this raises the probability of investment for any level of expected benefits. Second, however, the lower startup cost also reduces the value of investing today because it is now less expensive to wait and invest in the future. More precisely, the lower startup cost in future periods will result in an increase in  $EV(s_{it+1}|\omega_{it}, rd_{it} = 0)$ , the expected future value if the firm does not choose to invest in R&D. This leads to a reduction in  $\Delta EV$ , which reduces the firm's incentive to invest. Table 13 shows that the estimated reduction in  $\Delta EV$  is 0.455 million euros after five years of cost reduction in the high-tech industry and 0.104 million euros in low-tech. The effect of the startup cost reduction on the expected benefit of investing causes some firms to stop their R&D investments. Overall, the startup cost reduction has a less powerful effect on investment incentives than the reduction in maintenance cost.<sup>24</sup> However, it is important to point out that the two cost changes are not equivalent in terms of the overall cost of the subsidy. The maintenance cost reduction is applicable to all investing firms, whereas the startup cost reduction only applies to the firms that begin to invest in R&D.

## 8. Conclusions

■ Much of the empirical innovation and productivity literature focuses on measuring the private return to R&D investment, with the knowledge production function model being used as the primary framework. In this production model, firm R&D investments accumulate and depreciate over time, creating a knowledge stock that enters as an additional input into the production function. Estimates of the marginal product of this knowledge stock provide a measure of the return to R&D.

In this article, we take a different approach to measure the expected long-run, private payoff to R&D investment by estimating a model of the firm's dynamic decision to invest in R&D. The model allows the firm's R&D investment to raise the probability of being on a higher future productivity and profit path. Investment payoffs are realized in the future and subject to several sources of uncertainty. The proportional difference in the expected value of the firm between these two paths, therefore, is a crucial component of the firm's R&D decision and provides a natural measure of the expected payoff to the R&D investment.

The empirical model is designed to exploit the microdata collected in the Community Innovation Surveys. For Germany, this includes firm panel data on R&D activity, their realized product and process innovations, and variables to construct productivity and short-run profits. The four key structural components of the model are: the firm's profit function that relates productivity to profit, the evolution of firm productivity, which depends on product and process innovations realized by the firm, the probability of a product or process innovation given the firm's R&D choice, and the costs the firm must incur to generate an innovation.

The structural parameter estimates show firms that invest in R&D have a higher probability of realizing a product or process innovation, but R&D investment is neither necessary nor sufficient

<sup>24</sup> In contrast to this finding, Gonzalez, Jaumandreu, and Pazo (2005) find that subsidies would have a substantial impact on the R&D investment of Spanish manufacturing firms. They estimate that among firms that did not invest in R&D, half of the large ones would begin investing with a 10% cost subsidy and one third of the small ones would begin with a 40% subsidy.

for firm innovation. On average, across the high-tech manufacturing industries, the probability a firm has either a product or process innovation, given that they do not invest in R&D, is 0.229 but increases to 0.897 if they choose to invest. Firms in the low-tech industries have a lower probability of realizing innovations. On average, the probability is 0.208 if they do not invest in R&D and 0.804 if they do. Second, product innovation and process innovation lead to increases in future firm productivity, but product innovations are more important in the high-tech industries, raising productivity by 3.6%, whereas process innovations are more important for firms in the low-tech industries, raising their productivity by 3.5%. Third, firm productivity is persistent over time, hence, innovations that raise productivity have long-lived effects on firm value. Fourth, the cost to generate an innovation is significantly smaller for small firms and firms that have prior R&D investment.

We find that the expected payoff to investing in R&D varies across industries, averaging 4.4 million euros for the high-tech industries and 0.78 million euros for the low-tech industries. The payoff also varies substantially across firms within an industry with differences in age, capital stock, and productivity. Older, larger, and more productive firms have higher expected benefits from R&D investment, with productivity differences generating the most substantial differences across firms. We measure the expected long-run gain to investing in R&D as the proportional impact on firm value. In the five high-tech industries, the median gain across firms varies from 0.057 in the instrument industry to 0.085 in the vehicle industry, and averages 0.067 across the five industries. Across the seven low-tech industries, the median gain varies from 0.017 to 0.039 and averages 0.028.

Our structural model of firm R&D investment provides a decision rule for firm investment choice, allowing us to conduct counterfactual experiments to study the effect of changes in the economic environment on the R&D decision and future productivity. Additionally, we simulate different subsidy policies by changing the cost of innovation. The effect of lower innovation costs is at the heart of many policy discussions regarding the costs and benefits of public subsidies for R&D investment. The results show that in the high-tech industries, a 20% reduction in the innovation cost for an experienced R&D firm leads, after 10 years, to an average increase of 9.0 percentage points in the probability of investing in R&D, a 1.40% increase in mean firm productivity, and an increase in the expected benefit of R&D of 0.5 million euros. The same proportional reduction in the cost faced by high-tech firms that start to invest in R&D has virtually no impact on the probability of investing or the level of productivity. In contrast, in the low-tech industries, where there is a lower overall investment rate by firms, the startup margin is more important. A 20% reduction in the startup cost raises the investment rate by 1.2 percentage points and mean productivity by 7.0% after 10 years. The simulations illustrate that the effect of innovation cost reduction on R&D decisions and productivity growth depends on industry, the distribution of expected benefits, and, importantly, the target group of the subsidy. A structural model of the long-run payoff to R&D is needed to assess the impact of cost subsidies on the firms' investment decisions.

## References

- AW, B.Y., ROBERTS, M.J., AND XU, D.Y. "R&D Investment, Exporting and Productivity Dynamics." *American Economic Review*, Vol. 101 (2011), pp. 1312–1344.
- BERNSTEIN, J.I. AND NADIRI, M.I. "Research and Development and Inter-Industry Spillovers: An Empirical Application of Dynamic Duality." *Review of Economic Studies*, Vol. 56 (1989), pp. 249–269.
- . "Product Demand, Cost of Production, Spillovers, and the Social Rate of Return to R&D." NBER Working Paper no. w3625, 1991.
- COHEN, W.M. "Fifty Years of Empirical Studies of Innovative Activity and Performance." In B.H. Hall and N. Rosenberg, eds., *Handbook of the Economics of Innovation*. Vol. 2, Chapter 4. North-Holland, Amsterdam: Elsevier, 2010.
- CRÉPON, B., DUGUET, E., AND MAIRESSE, J. "Research Innovation and Productivity: An Econometric Analysis at the Firm Level." *Economics of Innovation and New Technology*, Vol. 7 (1998), pp. 115–158.

- CZARNITZKI, D., HALL, B.H., AND ORIANI, R. "The Market Valuation of Knowledge Assets in US and European Firms." In D. Bosworth and E. Webster, eds., *The Management of Intellectual Property*. Cheltenham, UK: Edward Elgar, 2006.
- DORASZELSKI, U. AND JAUMANDREU, J. "R&D and Productivity: Estimating Endogenous Productivity." *Review of Economic Studies*, Vol. 80 (2013), pp. 1338–1383.
- EUROSTAT. *Nace Rev 1 - Statistical classification of economic activities in the European Community, Luxembourg*, 1996.
- FOSTER, L., HALTIWANGER, J., AND SYVERSON, C. "Reallocation, Firm Turnover and Efficiency: Selection on Productivity or Profitability." *American Economic Review*, Vol. 98 (2008), pp. 394–425.
- GONZALEZ, X., JAUMANDREU, J., AND PAZO, C. "Barriers to Innovation and Subsidy Effectiveness." *RAND Journal of Economics*, Vol. 36 (2005), pp. 930–950.
- GRILICHES, Z. "Issues in Assessing the Contribution of Research and Development to Productivity Growth." *Bell Journal of Economics*, Vol. 10 (1979), pp. 92–116.
- . "The Search for R&D Spillovers." *The Scandinavian Journal of Economics*, Vol. 94 (1992), pp. 29–47.
- . "R&D and Productivity: The Unfinished Business." In *R&D and Productivity: The Econometric Evidence*. Chapter 12. Chicago: The University of Chicago Press, 1998.
- HALL, B.H. "The Private and Social Returns to Research and Development." In B.L.R. Smith and C.E. Barfield, eds., *Technology, R&D, and the Economy*. Washington, D.C.: The Brookings Institution, 1996.
- . "Innovation and Productivity." NBER Working Paper no. w17178, 2011.
- HALL, B.H., MAIRESSE, J., AND MOHNEN, P. "Measuring the Returns to R&D." In B.H. Hall and N. Rosenberg, eds., *Handbook of the Economics of Innovation*. Vol. 2, Chapter 22. North-Holland, Amsterdam: Elsevier, 2010.
- HASHMI, A.R. AND VAN BIESEBROECK, J. "The Relationship between Market Structure and Innovation and Market Equilibrium: A Case Study of the Global Automobile Industry." *Review of Economics and Statistics*, Vol. 98 (2016), pp. 192–208.
- KLETTE, T.J. "R&D, Scope Economies, and Plant Performance." *RAND Journal of Economics*, Vol. 27 (1996), pp. 502–522.
- KLETTE, T.J., MOEN, J., AND GRILICHES, Z. "Do Subsidies to Commercial R&D Reduce Market Failures? Microeconomic Evaluation Studies." *Research Policy*, Vol. 29 (2000), pp. 471–495.
- MAIRESSE, J. AND MOHNEN, P. "Using Innovation Surveys for Econometric Analysis." In B.H. Hall and N. Rosenberg, eds., *Handbook of the Economics of Innovation*. Vol. 2, Chapter 26. North-Holland, Amsterdam: Elsevier, 2010.
- MAIRESSE, J., MOHNEN, P., AND KREMP, E. "The Importance of R&D and Innovation for Productivity: A Reexamination in Light of the French Innovation Survey." *Annales d'Economie et de Statistique*, No. 79–80 (2005), pp. 487–527.
- MOHNEN, P. AND HALL, B.H. "Innovation and Productivity: An Update." *Eurasian Business Review*, Vol. 3 (2013), pp. 47–65.
- OECD. *Oslo Manual: Proposed Guidelines for Collecting and Interpreting Technological Innovation Data*. 1st, 2d, and 3d ed, Paris, 1992, 1997, 2005.
- OLLEY, G.S. AND PAKES, A. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, Vol. 64 (1996), pp. 1263–1297.
- PETERS, B., ROBERTS, M.J., AND VUONG, V.A. "Dynamic R&D Choice and the Impact of the Firm's Financial Strength." *Economics of Innovation and New Technology*, Vol. 26 (2017), pp. 134–149.
- RAMMER, C. AND PETERS, B. "Innovation Panel Surveys in Germany." In F. Gault, ed., *Handbook of Innovation Indicators and Measurement*. Cheltenham, UK and Northampton, MA, USA: Edward Elgar, 2013.
- ROBERTS, M.J. AND VUONG, V.A. "Empirical Modeling of R&D Demand in a Dynamic Framework." *Applied Economic Perspectives and Policy*, Vol. 35 (2013), pp. 185–205.
- ROBERTS, M.J., XU, D.Y., FAN, X., AND ZHANG, S. "A Structural Model of Demand, Cost and Export Market Selection for Chinese Footwear Producers." NBER Working Paper no. w17725, 2012.
- ROGERS, M. "R&D and Productivity: Using UK Firm-Level Data to Inform Policy." *Empirica*, Vol. 37 (2010), pp. 329–359.
- RUST, J. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*, Vol. 55 (1987), pp. 999–1033.
- XU, D.Y. "A Structural Empirical Model of R&D Investment, Firm Heterogeneity and Industry Evolution." Working Paper, Duke University, 2008. <https://sites.google.com/site/yixusite>. Accessed February 24, 2017.