A new approach to estimating the private returns to R&D

Ådne Cappelen* Pierre Mohnen† Arvid Raknerud‡
Marina Rybalka§

February 27, 2023

Abstract

This paper revisits the estimation of the private returns to R&D. In the standard approach, the returns to R&D only incorporates the impact of increased R&D on productivity or production costs. In contrast, we allow for endogeneity of output and variable inputs, where we only condition on the quasi-fixed inputs, R&D and physical capital stock. Our empirical analyses are based on an extended Cobb Douglas production function that allows for firms with zero R&D capital, which is useful for studying transitions from being R&D non–active to become R&D active. We further accommodate heterogeneity in R&D intensities, where rates of return may vary over a firm's life cycle, which is particularly relevant to R&D starters. Using a comprehensive panel of Norwegian firms observed in the period 2001-2018, we estimate average private returns to be 7-8 percent, which is substantially lower than estimates commonly reported in the literature. The low returns could reflect a high public subsidy rate of R&D, and indicate that, without large public subsidies, the aggregate level of private R&D in Norway would have been much lower than the current level.

JEL classification: C33, C52, D24, O38
Keywords: returns to R&D, productivity, subsidies, R&D capital stock, dynamic panel data

*Statistics Norway, Research Department, E-mail: adne.cappelen@ssb.no
†Maastricht University and UNU Maastricht Economic and Social Research Institute on Innovation and Technology (UNU-MERIT), E-mail: p.mohnen@maastrichtuniversity.nl
‡Statistics Norway, Research Department, E-mail: rak@ssb.no
§Statistics Norway, Department of Firm Dynamics, E-mail: ryb@ssb.no

1 Introduction

The private returns to R&D are a parameter of key interest to many public and private agents, such as investors, managers and policy makers. It is often found to be higher than for other investments, which is used by interest groups and others to propose policies to support private R&D, including broad R&D tax credits schemes. Even if some schemes may generate little in terms of spillover effects (see Nilsen et al., 2021), evidence of a high general private returns to R&D might convince the public that policies to support R&D are worth their costs to taxpayers. To assess the private returns to R&D as accurate as possible is important as public R&D support are huge in most developed countries. For example, in Norway total public R&D support (tax credits plus grants) to business enterprise R&D (BERD) increased from 0.11 to 0.22 percent of GDP from 2006 to 2018, while BERD increased from 0.78 to 1.05 percent. In comparison, the average public support of BERD in the OECD area increased from 0.15 to 0.18 percent of GDP in the same period (OECD, 2022)

While there are many approaches to estimating the returns to R&D in the literature – whether primal or dual (see the survey by Hall et al., 2010) – they almost all have in common that they derive returns estimates from the productivity impact of R&D under a *ceteris paribus* condition. In the *primal approach*, the marginal returns (R) to R&D means the increase in output (Y) as a result of an increase in R&D (F). For example, Doraszelski and Jaumandreu (2013) refers to the returns to R&D as the elasticity of output with respect to R&D expenditures. However, it is more common to transform elasticity estimates (β) into marginal returns estimates from the relation: $\beta = R(F/Y)$ – which is simply the definition of R&D elasticity. In the *dual approach*, R refers to the corresponding reduction in costs for given output. For example, Bernstein and Nadiri (1988) defines R as the real cost reduction from a unit increase in F for given Y.

By far, the most common way of specifying the underlying production function in the literature is to use a Cobb–Douglas function with equal elasticities across firms. However, according to R&D surveys, most firms report that they do not undertake any R&D, which would imply infinite returns at the extensive margin. The standard solution is to estimate the average returns using only firms with positive R&D, e.g.

excluding firm-year observations before a firm becomes R&D active. This creates a sample selection problem that may bias the results. Moreover, it does not remove the problem that returns will tend to be very high at the margin for firms with low R&D intensity, which is likely to inflate estimates of average returns purely because of a functional form assumption. A more flexible assumption might be that firms have different elasticities with respect to R&D, which is consistent with the huge heterogeneity in R&D intensity observed in the data.

A second problem with the interpretation of returns as simply meaning productivity is that firms do not change their input of R&D *ceteris paribus*. The effect of partially increasing R&D should arguably also incorporate the effects of adjusting other inputs in response to increased R&D, most obviously intermediate inputs and labour. A value added function, or even profit function, might be better suited to assess the private returns to R&D than productivity. The profit function would capture both the direct and indirect effects on profits triggered by a (partial) change in R&D, whereas the value added function would capture both increased profits to owners and the increased earnings to employees.

We address the limitations of the existing literature in three ways. First, we propose an extended Cobb–Douglas production function, where output depends on a *translation* of the R&D capital stock, with an unknown but estimable translation parameter, to allow positive output from firms with zero R&D. This functional form may be particularly useful to analyze R&D starters, i.e. the transition from being a R&D non-performer to being R&D active. Secondly, we accommodate heterogeneity in observed R&D intensities by allowing the R&D elasticity to be firm-specific. This would explain the huge heterogeneity in R&D intensities across firms and why most firms do not engage in R&D at all. Third, we derive returns estimates from a value added function, where firms simultaneously optimize material and labour inputs for any level of the R&D capital stock – as opposed to a returns measure that only incorporates the (first order) impact of increased R&D on output or production costs.

In this study, we analyze a panel of Norwegian firms in all industries from 2001 to 2018. According to our preferred specification, our estimate of the weighted average gross rate of return to R&D spending by Norwegian firms is about 7-8 percent, using the amounts of R&D investment as weights (henceforth referred to as *R&D-weighted average return*). This estimate is low compared to the rate of return commonly observed in the international literature, cf. Hall et al. (2010).

The structure of the paper is as follows. Section 2 presents some studies relevant to our investigation. Section 3 describes our theoretical framework for analyzing the private returns to R&D, Section 4 presents the econometric model, Section 5 the data, Section 6 the results and Section 7 offers some concluding comments.

2 Approaches to studying the relation between R&D and productivity

Several models of the relationship between R&D investment and productivity at the firm level have been proposed in the empirical literature. One general model structure, usually referred to as the CDM model, which was proposed by Crepon, Duguet and Mairesse (1998) following a conceptual model by Pakes and Griliches (1998). Here firm output is a function of input services and total factor productivity. Under the assumption of a standard neoclassical production function, output (Y) can be expressed as a function of the input factors labour (L), intermediates (M) and capital (K):

$$Y = A^* f(L, M, K). \tag{1}$$

where A^* is referred to as total factor productivity, which is assumed to depend on several variables relating to R&D, market factors, industry, and possibly other variables. One way of refining this model is to include additional variables in Equation (1) to capture the effect of intangible investments – both internal and external to the firm. One such factor is R&D, which is not directly treated as primary input in the CDM framework, but is instead assumed to influence A^* in equation (1) through product and process innovations.

There is also a long tradition in economics of specifying R&D as a primary input factor, F, in a standard Cobb-Douglas production – rather than indirectly via product- and process innovations as in the CDM model. The factor F is then often treated as an intangible capital stock (R&D capital), which similar to physical capital, K, accumulates by means of investments according to the PIM method:

$$F_t = (1 - \delta)F_{t-1} + I_{t-1}, \tag{2}$$

where δ is the depreciation rate of the R&D capital stock and I (real) R&D investment. Thus, if an estimate (or qualified guess) of the depreciation rate is available, one can calculate the R&D capital stock, F, using (2), and estimate the production function (1) directly, which then can be used to estimate the returns to R&D investments, I.

In recent years, with access to more micro data, more attention has been devoted to the role of intangible capital, such as information and communication technology (ICT) and organizational capital (see, for instance, Hall et al. 2013; Acemoglu and Restrepo 2018, 2020; Autor and Salomons 2018; Leduc and Liu 2019). Increased quality of workers is also an important source of productivity growth. For example Piekkola (2020) makes a distinction between R&D labour, management and advertising, and general labour, and constructs a measure of labour input quality based on the share of employees in each category and their relative wages. Our approach is broadly in line with Piekkola (2020), with the distinction that we measure labour quality by educational attainments rather than by dividing employees into professional categories.

An important feature of the (standard) Cobb-Douglas production function framework is that it cannot be applied to all firms without modifications, as it predicts zero output for firms with zero R&D. In the literature using micro data, there are several options available to circumvent that problem. One "solution" is simply to study those firms that report positive R&D and neglect other firms. This strategy does not solve the problem with firms that become R&D active during the observation period. Including these firms from that year on creates a sample selection problem that may bias estimates of the returns to R&D at the extensive margin. Since *potential* R&D performers are often a target of public R&D policies, the bias is potentially important. The problem of sample selection can be solved ad hoc by adding a small amount of R&D investment to firms with zero reported R&D, which makes it technically possible to include them in the analysis. A refinement of this solution is suggested by Griffith et al. (2006) and Hall et al. (2013). Relying on the CDM approach, they replace observed R&D spending with imputed R&D using data for all firms. In this way, zero R&D investment is replaced by nonzero imputed R&D. While this approach may perhaps be justified for firms who report zero R&D in *some* years, it does not not allow us to study the returns to R&D on the extensive margin, i.e. of *becoming* R&D active. Finally, one may specify a more flexible functional form that allows zero R&D, as suggested already by Griliches (1979). The advantage of this solution, which we will pursue here, is that one avoids altering the data or the sample taking all observations at face value.

3 Theoretical framework

Our starting point is an extended Cobb Douglas production function with labour, intermediates, tangible capital and R&D capital as inputs. The first extension is that we assume that the production function is homogenous of degree β in a *translation*, $\lambda + F$, of the R&D capital stock, F, thereby bounding the marginal returns to R&D and allowing positive output from firms with zero R&D. The second extension is that the production function is homogenous of degree ε in an aggregate function $g(L)$ of the vector $L = (L^{(1)}, L^{(2)}, L^{(3)})$ of man-years from three skill classes. Although there are examples of studies that control for the quality of labour input (e.g. Doraszelski and Jaumandreu, 2013), our appraoch is among the most elaborate in this respect (Doraszelski and Jaumandreu, *op. cit.*, only separate between temporary and permanent employees). Since R&D active firms generally hire more educated and higher paid workers than other firms, the assumption of homogeneous labour across education groups risks confounding the productivity effect of R&D with that of skill. Under the standard assumptions that the production function is homogenous with respect to physical capital and intermediates (of degree γ and ρ, respectively), we can write

$$Y = Ag(L)^{\varepsilon} M^{\rho} (\lambda + F)^{\beta} K^{\gamma} \quad (3)$$

where A is total factor productivity (unexplained "neutral efficiency"). Moreover, sales revenue equals $S = PY$, where P is the (potentially endogenous) output price, and value added equals $V = S - q_M M$, where q_M is the price of intermediates.

Importantly, the specification (3) allows the R&D capital stock, F, to be zero without implying $Y = 0$. In particular, the marginal output of R&D is:

$$Y'_F = \beta \frac{Y}{\lambda + F}$$

and the R&D elasticity is:

$$\mathrm{El}_F Y = \beta \frac{F}{\lambda + F}$$

An important property of (3) is that Y'_F, does not increase towards infinity as F tends to zero. When $\lambda = 0$, we have the Cobb-Douglas case with $\mathrm{El}_F Y = \beta$ and $Y'_K \to \infty$ as $F \to 0$. In this case, it will always be profitable to invest in R&D as the returns to R&D at the extensive margin is infinity.

In the empirical model, we assume that ε, ρ and γ are common parameters, but allow β to be firm-specific, as will be motivated by the analysis below.

3.1 Economic behavior

We assume that producers are price takers in all factor markets, but not in the product markets, and that both types of capital, K and F, are fixed in the short run, so that the short-run optimization of the firm is w.r.t. L and M for *pre-determined* R&D capital stock, F_{it}, and tangible capital stock, K_{it}. The corresponding labour cost function, i.e. given the level of aggregate labour input, $g_i(L)$, is

$$C_{it}(\mathbf{q}_{it}, g_i(L_{it})) = c_{it} g_i(L_{it}) \tag{4}$$

where $\mathbf{q}_{it} = (q_{it}^{(1)}, q_{it}^{(2)}, q_{it}^{(3)})$ is the vector of firm-specific wage rates of low-, medium- and high-skilled labor, respectively, and c_{it} is the firm-specific unit price of labour (the aggregate wage rate). Appendix A derives the formulas in (4) in the case of a CES aggregation function of labour inputs, L. We also allow $g_i(\cdot)$ to be firm-specific to be consistent with firms choosing $L_{it}^{(m)} = 0$ from some m, for example, only employing workers in the lowest skill category ($m = 1$).

We next consider the partial optimization problem of firm i in the beginning of period t conditional on the *predetermined* variables F_{it} and K_{it}, assuming that the firm knows $\mathbf{q}_{it}$, q_{Mt} and A_{it}. The problem is to choose the price that maximizes operating profits. Making the usual assumption of monopolistic competition with demand given by:

$$P_{it} = \Phi_{it} Y_{it}^{-e}$$

where e ($e > 1$) is the elasticity of demand with respect to sales price , P_{it} profit maximization gives the following equation for log value added:

$$\ln V_{it} = \widetilde{\theta} - \widetilde{\varepsilon} \ln c_{it} + \widetilde{\beta}_i \ln r_{it}(\lambda) + \widetilde{\gamma} \ln K_{it} - \widetilde{\rho} \ln q_{Mt} + \widetilde{a}_{it} \tag{5}$$

where $\widetilde{\theta}$ is a constant, $r_{it}(\lambda) = \lambda + F_{it}$, $\widetilde{\varepsilon} = \varepsilon\vartheta$, $\widetilde{\gamma} = \gamma\vartheta$, $\widetilde{\beta}_i = \beta_i\vartheta$, $\widetilde{\rho} = \rho\vartheta$, and $\widetilde{a}_{it} = \vartheta(\ln A_{it} + \ln \Phi_{it}/(e-1))$, with

$$\vartheta = \frac{(e-1)}{(\varepsilon + \rho + e - e(\varepsilon + \rho))} \in (0, (1 - \varepsilon - \rho)^{-1}). \tag{6}$$

See Appendix A for a proof.[1] The left and right limit correspond to, respectively, $e \to 1$ and $e \to \infty$. Equation (5) will be the key equation for estimating the private returns to R&D. We show in Appendix A that the profits, Π_{it} (value added minus wage costs), is proportional to V_{it}:

$$\frac{\Pi_{it}}{V_{it}} = \frac{e - (e-1)(\rho + \varepsilon)}{e - (e-1)\rho} \in (\frac{1 - \rho - \varepsilon}{1 - \rho}, 1)$$

where the left and right limit correspond to, respectively, $e \to \infty$ and $e \to 1$.

We have no information about firm-specific intermediate input prices, therefore the term involving $\ln(q_{Mt})$ in Equation (5) cannot be distinguished from time dummies in our empirical specification and $\widetilde{\rho}$ cannot be identified. On the other hand, we do observe firm-specific wages. Although the aggregate wage rate c_{it} is in general an unknown function of $\mathbf{q}_{it}$, this problem can be overcome using the *Sato-Vartia* index:

$$\frac{c_{it}}{c_{i,t-1}} = \prod_k \left(\frac{q_{it}^{(k)}}{q_{i,t-1}^{(k)}} \right)^{\omega_{it}^{(k)}}$$

where the weights, $\omega_{it}^{(k)}$, are proportional to the geometric average of the (observable) cost shares $\alpha_{it}^{(k)}$ and $\alpha_{i,t-1}^{(k)}$ of skill class k:

$$\alpha_{it}^{(k)} = \frac{q_{it}^{(k)} L_{it}^{(k)}}{\sum_{k=1}^{3} q_{it}^{(k)} L_{it}^{(k)}} \tag{7}$$

The Sato-Vartia index is exact in the case of the CES aggregator function where $g_i(L) = g(L; \mathbf{a}_i)$ for weight parameters $\mathbf{a}_i$ (see Appendix A for formulas). In that

[1] $\widetilde{\theta}$ is defined in Equation (26) in Appendix A

case we have the well-known result that $\alpha_{it}^{(k)} = a_i^{(k)} \left(q_{it}^{(k)} / c_{it} \right)^{1-\sigma}$ (see Diewert 1978 and Brasch et al. 2022). In the special case of Cobb-Douglas ($\sigma = 1$), we get the famous result that factor cost shares are constant: $\alpha_{it}^{(k)} = a_i^{(k)}$. More important for our purpose is that the Sato-Vartia index is consistent to second order with any exact, twice differentiable aggregator function $g_i(\cdot)$. We can therefore apply the Sato-Varita index to obtain:

$$\Delta \ln c_{it} = \sum_k \omega_{it}^{(k)} \Delta \ln q_{it}^{(k)} \tag{8}$$

with the normalizing initial condition $c_{i1} = 1$.

3.2 The returns to R&D

To simplify notation, define

$$R_{it} = \frac{dV_{it}}{dF_{it}} = \frac{\widetilde{\beta}_i V_{it}}{F_{it} + \lambda} \tag{9}$$

which is our proposed value added-based measure of the private returns to R&D, as motivated in the Introduction. In the tradition of Hall et al. (2010), it is often assumed that R_{it} varies randomly about a common mean, R, where R is the constant marginal *cost* of R&D. To apply this assumption in our context, where F and K are pre-determined – and therefore based on ex ante expected returns – we define $V_{it}^e = E(V_{it} \, | F_{it}, K_{it})$ and assume the existence of a steady state defined as follows:

$$E(R_{it} | F_{it}, K_{it}) = \frac{\widetilde{\beta}_i V_{it}^e}{F_{it} + \lambda} = R \tag{10}$$

The first equality follows from (9), assuming

$$V_{it} = V_{it}^e + e_{it}$$

for an error term, e_{it}, whereas the second equality says that in a steady state with $F_{it} > 0$, expected returns equals the marginal costs of R&D. The interpretation of Equation (10) is that of an equilibrium correction, where a firm over time adjusts F_{it} towards a *firm-specific* equilibrium R&D intensity (but not necessarily R&D level). As we will discuss below, the adjustment may be sluggish and hampered by adjustment cost and uncertainty, so that in general $R_{it} \neq R$.

Under the above formal assumptions:

$$\widetilde{\beta}_i V_{it} = (F_{it} + \lambda)R + \widetilde{\beta}_i e_{it} \tag{11}$$

Taking the expectation conditional on $F_{it} > 0$ on both sides of (11) yields:

$$\widetilde{\beta}_i = R f_i(\lambda) \tag{12}$$

where

$$f_i(\lambda) = \frac{E(F_{it}|F_{it} > 0) + \lambda}{E(V_{it}|F_{it} > 0)} \tag{13}$$

The function $f_i(\lambda)$, which varies across i because β_i does so, represents an equilibrium relation between F_{it} and V_{it}. We will refer to (12)-(13) as the constant marginal costs (CMC) model. In particular, the CMC model explains why some firms never engage in R&D activities: If $\widetilde{\beta}_i < R\lambda/V_{it}^e$ the marginal returns to R&D are always lower than R.

A natural estimator of R would an R&D weighted average of R_{it}, to ensure that equal weight is given to each NOK of investment. This approach would require that $\widetilde{\beta}_i$ is known or estimable. Since $\widetilde{\beta}_i$ is firm-specific, it cannot be estimated directly because of the incidental parameter problem. An alternative strategy is the following: (i) estimate $f_i(\lambda)$ using the empirical counterpart:

$$\widehat{f}_i(\lambda) = \frac{\sum_{t=1}^{T_i} 1(F_{it} > 0)(F_{it} + \lambda)}{\sum_{t=1}^{T_i} 1(F_{it} > 0)V_{it}} \tag{14}$$

where $1(A)$ denotes the indicator function which is one if the statement A is true, and (ii) substitute the incidental parameters $\widetilde{\beta}_i$ using (12) (replacing $f_i(\lambda)$ with $\widehat{f}_i(\lambda)$) and estimate R using GMM (see Section 5 for details).

In the literature, the usual assumption is that $\beta_i = \beta$ (no heterogeneity in the elasticity of Y w.r.t. F), implying $f_i(\lambda) = f(\lambda)$ for all i under the CMC assumption. We will refer to this special case as the restricted CMC model (R-CMC), which can be stated as:

$$\widetilde{\beta} = R f(\lambda) \tag{15}$$

The function $f(\lambda)$ represents an equilibrium R&D intensity that does *not* depend on i because β is assumed to be homogeneous across firms. According to the H-CMC model, there should be no systematic or persistent differences in R&D intensities across firms. To estimate (weighted) average returns from the R-CMC-model, we can do the following: (i) estimate the common parameter $\widetilde{\beta}$ using GMM, (ii) obtain individual R_{it} estimates from Equation (9), and (iii) calculate the (weighted) average of R_{it} across (i, t).

By definition, $f_i(\lambda)$ refers to an equilibrium state where the investment ratio $i_{i,t-1} = I_{i,t-1}/F_{it}$ is close to δ on average. In that case, $R_{it} \simeq R$. In contrast, $i_{i,t-1} = 1$ for *R&D starters.* More generally, it is reasonable to expect that firms for which a large proportion of the $i_{i,t-1}$ observations are either close to one (say, above some value α_{high}) or close to zero (say, below α_{low}) are far away from their equilibrium R&D intensity. In that case $\widehat{f}_i(\lambda)$ will be biased as an estimator $f_i(\lambda)$. A simple way to capture this is to assume that:

$$f_i(\lambda) \simeq \widehat{f}_i(\lambda) \left(1 + \lambda_{low} 1(\bar{i}_i \leq \alpha_{low}) + \lambda_{high} 1(\bar{i}_i \geq \alpha_{high})\right) \tag{16}$$

where λ_{low} and λ_{high} represent average relative estimation biases and

$$\bar{i}_i = \frac{\sum_{t=1}^{T_i} I_{i,t-1}}{\sum_{t=1}^{T_i} F_{i,t}} \tag{17}$$

is the average investment rate of the firm. For example, if $\widetilde{\beta}_i \simeq R\widehat{f}_i(\lambda)(1 + \lambda_{high})$ a positive estimate of λ_{high} would indicate that the equilibrium R&D intensity of firms with very high investment rates are underestimated. From (9), the weighted average return is:

$$\sum_{t=1}^{T_i} \omega_{it} R_{it} \simeq R(1 + \lambda_{high}) > R$$

where $\omega_{it} = (F_{it} + \lambda)/\sum_{t=1}^{T_i}(F_{it} + \lambda)$ is the ideal R&D weight. Thus the parameter λ_{high} would capture a high observed return to R&D in firms that are well below their equilibrium R&D intensity. This is exactly what we would expect from R&D starters. In contrast, a negative λ_{low} could reflect low returns to R&D for firms above their equilibrium R&D intensity.

It would not be unusual for firms to have some extreme investment rates. For example, it is well known that in the presence of non-convex adjustment costs of capital, there will be an "interval of inaction", where the firm will not decrease/increase F_{it} even if the equilibrium condition (10) is violated. This could be the case if the selling price of capital is lower than the purchasing price (full or partial irreversibility) or there are fixed costs of investments (see for example Golombek and Raknerud, 2018). A recent, but sparse, literature on the implications of fixed costs on the time series properties of the returns to R&D, also suggests that fixed costs will cause higher rates of return to firms that invest relatively more in R&D, i.e. have a high $I_{i,t-1}/F_{it}$ (see Resutek 2022).

4 Sample and variable construction

For our analysis, we have constructed a panel of annual firm-level data for Norwegian firms with at least three consecutive observations during 2001–2018. The base for the sample is the R&D statistics, which are survey data collected by Statistics Norway. These data comprise detailed information on firms' R&D activities, such as total R&D expenses (divided into internally performed R&D and externally purchased R&D), the number of employees engaged in R&D activities, and the number of man-hours worked in R&D. Only firms with more than 50 employees are automatically included in the survey. For smaller firms (with 5–49 employees) a stratified sampling scheme is employed. The stratification is based on industry classification (NACE codes) and firm size. However, these smaller firms are not representative of firms of their size and industry, because they have a higher probability of engaging in R&D. Currently, data are available for 1993, 1995, 1997, 1999, and *annually* from 2001 to 2018. To supplement the regular R&D census, we obtained questionnaire data from SKF on each applicant's R&D expenditure for three years prior to applying for tax credits. These data are collected by the RCN, which must approve in advance any project that is to form the basis for tax credits. The information from all available surveys is used for the construction of R&D capital stocks.

The survey data on R&D are supplemented with data from three different registers: The accounts statistics, the Register of Employers and Employees (REE), and the National Education Database (NED) and the R&D Support Database[2]. The latter contains information about each firm's R&D support during 2001-2019 – both direct support and tax credits. There are only a few examples in the data of

[2]In Norwegian: Virkemiddeldatabasen

firms that report positive R&D in the annual R&D survey and never obtains R&D tax credits. Descriptive statistics are provided in Appendix B.

Value added, V, is net value added at factor cost and computed as the sum of operating profits net of depreciation and labor costs. All prices have been deflated by the same price index so that in any time period, one dollar of any cost component has the same value as one dollar of a revenue component. We use the price index for R&D investment as a common deflator. This is based on the price indices from the national accounts for the various components making up total R&D, implying that the real unit price of R&D is one. According to Hall et al. (2010) the choice of deflator usually does not matter much for the econometric results for the main parameters of interest. R&D investment, I, is yearly R&D investment as they are reported in the survey data. The (real) R&D capital stock (F) at the beginning of a given year t is computed by the perpetual inventory method (2) using a constant rate of depreciation $\delta = 0.15$. The benchmark for the initial R&D capital stock for a given firm in its first observation year in 2001-2018 is based on its history of R&D investment going back to (at most) 1993 (for details, see Cappelen et al., 2012).

To construct the physical capital stock, K, we used information from the accounts statistics. The accounts statistics distinguish between several groups of physical assets. To obtain consistent definitions of asset categories over the whole sample period, all assets have been divided into only two types: equipment, denoted by e, which includes machinery, vehicles, tools, furniture and transport equipment, and buildings and land, denoted by b. The expected lifetimes of the physical assets in group e (of about 3–10 years) are considerably lower than those of the assets in group b (about 40–60 years). Total capital, K, is then an aggregate of equipment capital, e, and building capital, b. We use the book value as a measure of the capital stock. When aggregating the two capital types, we use a Törnqvist volume index with time-varying weights that are common across firms in the same industry (see OECD, 2001).

Man-hours, $L^{(k)}$, is the sum of all individual man-hours worked by employees in the given firm according to the contract. For each firm, we distinguish between three skill groups: employees with primary, secondary and postsecondary education (see Appendix B). Man-hours worked by persons in skill group k are aggregated to the firm level to construct $L^{(k)}$. When calculating the (average) wage in each skill group, q_{kit}, we use predicted wages from a wage regression with random individual effects, where we include a dummy for skill-category (k) and dummies for industry (NACE 2), region, gender and calendar year as regressors. It is the average of the *predicted* wages for all the firm's employees in the given skill category from this regression which is the basis for calculating q_{kit}. Using matched employer-employee data we are able to match each firm with its registered employees over time. This prediction-based method is chosen to reduce the problem of errors in reported hours in the employer-employee register. Errors are often related to part-time employees and/or timeliness problems, because hours are reported by week and wage costs by year. When estimating the wage equation we therefore restrict the sample to full time employees during the given calendar year.

5 Empirical analyses

Our dependent variable in the empirical analysis is $\ln V_{it}$. Our stochastic specification of the structural equation (5) is the following:

$$\ln V_{it} = -\widetilde{\varepsilon} \ln c_{it} + \widetilde{\gamma} \ln K_{it} + \widetilde{\beta}_i \ln r_{it}(\lambda) + a_i + \mu_t^* + \zeta_{it} \tag{18}$$

where a_i is a fixed effect, μ_t^* is the time-varying intercept (which incorporates the term $\widetilde{\rho} \ln q_{Mt}$), and ζ_{it} is the error term assumed to follow a first-order autoregressive process:

$$\zeta_{it} = \phi \zeta_{i,t-1} + e_{it} \tag{19}$$

with

$$|\phi| \leq 1, \ E[e_{it}] = 0, \ E[e_{it}^2] = \sigma_e^2$$

and

$$\text{Cov}[e_{it}, e_{jt}] = 0 \text{ if } t \neq s \text{ or } i \neq j.$$

Multiplying (18) by ϕ and quasi-differencing, yield:

$$\begin{aligned} \ln V_{it} = {} & \phi \ln V_{i,t-1} - \widetilde{\varepsilon} \ln c_{it} + \phi \widetilde{\varepsilon} \ln c_{i,t-1} + \widetilde{\beta}_i \ln r_{it}(\lambda) - \phi \widetilde{\beta}_i \ln r_{i,t-1}(\lambda) \\ & + \widetilde{\gamma} \ln K_{it} - \phi \widetilde{\gamma} \ln K_{i,t-1} + v_i + \mu_t + e_{it} \end{aligned} \tag{20}$$

where $\mu_t = \mu_t^* - \phi\mu_{t-1}^*$ and $v_i = (1-\phi)a_i$. Next, we difference to eliminate the fixed effect, v_i:

$$\Delta \ln V_{it} = \phi\Delta \ln V_{i,t-1} - \widetilde{\varepsilon}\Delta \ln c_{it} + \phi\widetilde{\varepsilon}\Delta \ln c_{i,t-1} + \widetilde{\beta}_i\Delta \ln r_{it}(\lambda) - \phi\widetilde{\beta}_i\Delta \ln r_{i,t-1}(\lambda) + \widetilde{\gamma}\Delta \ln K_{it} - \phi\widetilde{\gamma}\Delta \ln K_{i,t-1} + \Delta\mu_t + \Delta e_{it} \quad (21)$$

5.1 GMM Estimator

We use lagged levels of the endogenous variables as instruments for the endogenous variables in (21) and estimate the interest parameters using GMM – as proposed by Arellano and Bond (1991). However, this method requires that the (nuisance) parameter λ is known. To estimate *all* the parameters, we perform a grid search in λ-space to minimize the residual sum of squares associated with the GMM estimator for the given λ-value. Following the general methodology of Arellano and Bond (1991, the GMM-estimator uses the following moments:

$$E(\ln(V_{i,t-s})\Delta e_{it}) = 0$$
$$E(\ln(c_{i,t-s+1})\Delta e_{it}) = 0$$
$$E(\ln r_{i,t-s+1}(\lambda))\Delta e_{it}) = 0$$
$$E(\ln(K_{i,t-s+1})\Delta e_{it}) = 0$$

for $s \geq 2$ (see Equation(21)). That is, we treat *all* the right-hand side variables in equation (21) as pre-determined endogenous variables. A testable identifying assumption is that Δe_{it} is a MA(1) noise term. In principle, a non-linear GMM estimator could be applied, but, unfortunately, it runs into convergence problems as λ is poorly identified. Our results show that a very small λ minimizes the residual sum of squares, i.e. an estimate on par with the lowest positive F_{it} observation in our estimation sample.

As discussed in Section 3, we examine two specifications with respect to the marginal cost of R&D. The first is the CMC-model (12)-(13), where we assume a firm-specific $\widetilde{\beta}_i$ and a common costs of R&D (R) with a firm-specific steady-state R&D intensity, $f_i(\lambda)$. In this model, we substitute $\widetilde{\beta}_i$ with $Rf_i(\lambda)$ and replace $f_i(\lambda)$ with the right hand side of Equation (16), considering R as the GMM-estimand. The second case is a restricted version of this model, referred to as R-CMC, where we assume a common elasticity for all firms: $\widetilde{\beta}_i = \widetilde{\beta}$. In the R-CMC model, $\widetilde{\beta}$ (not R) is the estimand. It is straightforward to derive an estimate of R from $\widetilde{\beta}$ as a weighted average of R_{it}, as explained in Section 3.2.

5.2 Results

The parameter estimates are presented in Table 1. As a benchmark we also present a fixed-effects (FE) estimator of the R-CMC model. The FE estimator is the conventional within-estimator applied to equation (18). However, this method yields biased estimates in the presence of endogeneous *time-varying* explanatory variables, as described above.

Table 1: Estimates of coefficients in the value added equation under different model assumptions. Robust standard errors in parentheses

Indep. variables in structural equation*	Coeff.	GMM-estimates		FE-estimates
		CMC	R-CMC	R-CMC
		Est. (t-stat)	Est. (t-stat)	Est. (t-stat)
$\ln V_{i,t-1}$	ϕ	.228 (17.2)***	.260 (18.8)***	
$-\ln c_{it}$	$\widetilde{\varepsilon}$	.515 (2.82)***	.593 (2.74)***	.386 (3.88)***
$\ln K_{it}$	$\widetilde{\gamma}$	.090 (4.99)***	.107 (6.72)***	.126 (22.0)***
$\ln r_{it}(\lambda)$	$\widetilde{\beta}$		.022 (10.5)***	.013 (6.95)***
$\widehat{f}_i(\lambda)\ln r_{it}(\lambda)$	R	.063 (7.09)***		
$1(\bar{i}_i \leq \alpha_{low})\widehat{f}_i(\lambda)\ln r_{it}(\lambda)$	λ_{low}	.009 (0.29)		
$1(\bar{i}_i > \alpha_{high})\widehat{f}_i(\lambda)\ln r_{it}(\lambda)$	λ_{high}	.024 (3.43)***		
λ		.59	.59	.59
Number of obs.		33,105	33,105	39,070
Number of firms		4,524	4,524	4,719
R^2				0.22
Residual variance		.091	.096	.416

*$\widehat{f}_i(\lambda)$, α_{low} and α_{high} refers to Equations (12)-(14) and (16), where $[\alpha_{low}, \alpha_{high}]$ is the interquartile range of the investment rate, $\bar{i}_i$ defined in Equation(17), with $\alpha_{low} = 0.11$ and $\alpha_{high} = 0.29$.

Both the FE and GMM estimators of the coefficient related to the key R&D variable, $\ln r_{it}(\lambda)$, are positive and significant. In the case of the R-CMC model, this coefficient can be interpreted as the elasticity with respect to R&D (i.e., $\widetilde{\beta}$). The R-CMC specification is a special case of CMC where it is assumed that $\beta_i = \beta$ for all i. In the CMC model, the estimates related to $\widehat{f}_i(\lambda)\ln r_{it}(\lambda)$ can be interpreted in terms of average gross returns to R&D as a function of the firm-specific

Table 2: Distribution of marginal returns to R&D, R_{it}, conditional on $F_{it} > 0$ for different models and subsamples

Model (Estimator)	All obs.with $F_{it} > 0$	Subsample[1) Q[1]	Subsample[1) Q[2]-Q[3]	Subsample[1) Q[4]
CMC: heterogeneous elasticity (GMM)				
Weighted mean[2)]	.071	.072	.064	.122
Median	.073	.077	.064	.183
Unweighted mean	.383	.092	.082	1.71
R-CMC: common elasticity (GMM)				
Weighted mean	.081	.076	.075	.147
Median	.089	.104	.061	.221
Unweighted mean	2.48	4.04	1.40	3.31
R-CMC: common elasticity (FE)				
Weighted mean	.051	.051	.047	.084
Median	.066	.075	.046	.132
Unweighted mean	1.68	2.8	.935	2.02
Share of R&D (F_{it}) in sample	1	.212	.707	.080

Notes: [1)] $Q[n]$ refers to the n'th quartile (interval) of the R&D investment ratio, $\bar{i}_i$: $Q[1] = [0, .11)$, $Q[2] = [.11, .17), Q[3] = [.17, .29), Q[4] = [.29, 1]$
[2)]Weighted by share of R&D (F_{it}) in the given subsample

investment ratio, $\bar{i}_i$. We find that the estimated returns are highest on average in the fourth investment ratio quartile (about 2 percentage points higher than in the other quartiles combined). The equilibrium rate of return to R&D, R, is estimated to be 6.3 percent.

As expected, we find a significant positive relation between both types of capital and value added: the estimated long-run elasticity with respect to K lies in the range 0.09-0.13 using different estimation methods, and is of magnitude five times larger than the estimated (common) elasticity with respect to R&D capital, $\widetilde{\beta}$. The estimates of the autoregressive coefficient ϕ in Table 1 – the coefficient of $\ln V_{i,t-1}$ – lie in the range of 0.23-0.26 and are highly significant. Since the FE estimator uses the equations in levels (before quasi-differencing), ϕ is not identified using the FE estimator.

The coefficient of $-\ln(c_{it})$ is estimated to be about 0.5 in both the CMC and R-CMC model using GMM, and about 0.3 in the R-CMC model using the FE estimator. The results confirm that labour is the single most important production factor.

5.3 Returns to R&D

From the GMM estimates in Table 1, we can calculate the marginal return to R&D-investment, $R_{it} = \partial V_{it}/\partial F_{it}$, for each observation as explained in Section 3. The estimated marginal returns have a R&D-weighted mean of 7.1 and 8.1 percent in the CMC and R-CMC model, using GMM, and 5.1 in the R-CMC model, using the FE estimator (see Table 2). The low returns derived from the FE estimator reflect the low estimate of $\widetilde{\beta}$ (0.013) compared to using GMM (0.022). Mean and median values by quartiles ($Q[n]$) for the R&D-investment *ratio* ($I_{i,t-1}/F_{it}$) are also depicted. We see that the observations in Q[1] account for 21 percent of the total R&D stock, Q[2]-Q[3] (combined) account for 70 percent, and Q[4] account for only 8 percent. The latter category includes (almost all of) the R&D-starters during 2001-2018. As expected, the results show higher returns in the highest quartile, likely reflecting a disequilibrium phenomenon, as discussed in Section 3.2. However, there is no clear difference between Q[1] and Q[2]-Q[3]. In fact, we estimate slightly higher returns in Q[1] than on Q[2]-Q[3] combined. This is driven by the positive, but insignificant, λ_{low} estimate in Table 1.

Looking at the returns in the different investment ratio quartiles, and across models, we find some striking patterns. In the CMC model, average and median returns are more homogeneous across the quartiles than in the R-CMC model. In contrast, the R-CMC model depicts a pattern of much higher returns in Q[1] and Q[4] compared to the CMC model. These results indicate that the R-CMC model is too rigid. However, if we only care about the weighted average return to R&D, the estimates are remarkably similar – about 7-8 percent gross returns – both in the case of CMC and R-CMC using GMM.

The weighted average returns (7-8 per cent) are low compared to the rate of return commonly observed in the international literature , cf. Hall et al. (2010). Exceptions are Klette and Johansen (1998), who estimate the mean net rate of return to be 9 percent for Norwegian manufacturing, but with huge variations across industries and Parisi et al. (2006), who estimate the rate of return to knowledge capital to be only 4 percent on Italian data. The low return could reflect the high public subsidy rates mentioned in the Introduction, above 20 percent implied marginal subsidy rate in 2018 according to the OECD (2022). Our results indicate that without large public subsidies, the level of private R&D would need to be much lower to earn a normal rate of return on the investment.

5.4 Specification tests

As seen from Table 3, the Arellano–Bond test of zero first-order autocorrelation in the error term Δe_{it} is rejected, but not for second-order autocorrelation. This confirms that ζ_{it} follows a first-order autoregressive process, as assumed in (19). We also applied a Sargan test to test the validity of the overidentifying restrictions with regard to the instrumental variables. With a χ^2-test statistic of 894.2 and 805 degrees of freedom, we cannot reject this hypothesis. All these specification tests, seen together, give strong support to our econometric specification.

Table 3: Specification tests

	Observed value of test statistic (Z)	Level of significance $\Pr(Z > z)$
Test of zero autocorrelation in errors*		
order 1	-11.762	.001
order 2	-.35	.72
J-test of overidentifying restrictions**	894.2	.150

Notes: *t-test **test statistics is distributed as $\chi^2(805)$

6 Conclusions

This paper has revisited the estimation of the private return to R&D. We have proposed an extended Cobb Douglas production function which allows for firms with zero R&D capital in order to study the transition from being R&D non–active to active, without restricting the sample to R&D performers. In the standard approach, the returns to R&D only incorporates the impact of increased R&D on productivity or production costs. In contrast, we obtained returns estimates from a value added function derived under the assumption of profit maximizing firms that optimize labour and intermediate factor inputs at any level of R&D capital. The value added function captures both increased profits to owners and increased earnings to employees resulting from the R&D investment. We further accommodated the huge observed heterogeneity in R&D intensities by allowing R&D elasticities to be firm-specific, and incorporate heterogeneity in labour quality by distinguishing between three levels of educational attainments of employees. Estimating the model on a comprehensive panel of Norwegian firms observed in the period 2001-2018, we obtained estimates of the average private return of 7-8 percent, which is substantially lower than estimates commonly reported in the literature. The low return estimates could reflect a high public subsidy rate of R&D, and indicates that, without large public subsidies, the aggregate level of private R&D would have been much lower than the current level.

References

[1] Acemoglu, D. and P. Restrepo (2020): Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy.* 128 (6), 2188–2244.

[2] Arellano, M. and S. Bond (1991): Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies,* 58, 277–297.

[3] Autor, D. and A. Salomons (2018): Is automation labor displacing? Productivity growth, employment, and the labor share. Brookings Papers on Economic Activity 49 (1 (Spring), 1–87.

[4] Baily, M., C. Hulten and D. Campbell (1992): Productivity dynamics in manufacturing plants, *Brookings Papers on Economic Activity, Microeconomics*, Vol. 1992, 187–249.

[5] Brasch, T. von and A. Raknerud (2022): Productivity growth, firm turnover and new varieties. *Scand. J. of Economics*, 124, 646–676.

[6] Cappelen, Å., A. Raknerud and M. Rybalka (2012): The effects of R&D tax credits on patenting and innovations, *Research Policy*, 41, 334–345.

[7] Cohen, W. and D. Levinthal (1989): Innovation and learning: two faces of R&D, *Economic Journal*, 99, 569–596.

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

[9] Diewert, W.E. (1978): Superlative Index Numbers and Consistency in Aggregation, *Econometrica*, Vol. 46, No. 4 (July, 1978)

[10] Doraszelski, U. and J. Jaumandreu (2013): R&D and productivity: Estimating endogenous productivity, *Review of Economic Studies*, Vol. 80, No. 4, 1338-1383.

[11] Golombek, R. and A. Raknerud (2018): Exit dynamics of start-up firms. Structural estimation using indirect inference. *Journal of Econometrics*, 2018, 205, 204–225.

[12] Griffith, R., E. Huergo, J, Mairesse and B. Peters (2006): Innovation and Productivity Across Four European Countries, *Oxford Review of Economic Policy*, 22, 483-498.

[13] Griffith, R., S. Redding and J. Van Reenen (2004): Mapping the two faces of R&D: productivity growth in a panel of OECD countries, *Review of Economics and Statistics*, 86, 883–895.

[14] Griliches, Z. (1979): Issues in assessing the contribution of research and development to productivity growth, *Bell Journal of Economics*, 10, 92-116.

[15] Griliches, Z. (1995): R&D and productivity: econometric results and measurement issues, in P. Stoneman (ed.): *Handbook of the Economics of Innovation and Technical Change*, NBER and University of Chicago Press.

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

[17] Hall, B. H. (2002): The financing of research and development, *Oxford Review of Economic Policy*, 18, 35–51.

[18] Hall B. H., F. Lotti and J. Mairesse (2013): Evidence on the Impact of R&D and ICT investment on Innovation and productivity in Italian firms, NBER Working paper No. 18053.

[19] Hall, B. H., J. Mairesse and P. Mohnen (2010): Measuring the returns to R&D, in Chapter 24: *Handbook of the Economics of Innovation*, Vol. 2. North-Holland Publ. Comp., Amsterdam.

[20] Klette, T. J., and F. Johansen (1998): Accumulation of R&D capital and dynamic firm performance: A not-so-fixed effect model, *Annales de Economie et de Statistique*, 49/50, 389-419.

[21] Leduc, S. and Z. Liu (2019): Robots or Workers: A Macro Analysis of Automation and Labor Markets. Federal reserve Bank of San Francisco Working Paper Series (17).

[22] Mairesse, J. and P. Mohnen (2004): The importance of R&D for innovation: A reassessment using French survey data, *NBER Working Paper 10897*, National Bureau of Economic Research, Cambridge, MA.

[23] . Bernstein, J and M Ishaq Nadiri (1988):Interindustry R&D Spillovers, Rates of Return, and Production in High-Tech Industries. *American Economic Review*, vol. 78, issue 2, 429-34.

[24] Norges forskningsråd (2010): Det norske forsknings- og innovasjonssystemet - statistikk og indikatorer (The Norwegian research and innovation system – statistics and indicators), The Research Council of Norway, Oslo 2010.

[25] OECD (2001): *Measurement of Capital Stocks, Consumption of Fixed Capital and Capital Services*. OECD Manual 131, Paris, France, and Washington, DC.

[26] OECD (2022): *OECD R&D tax incentives database, 2021 edition.*

[27] Pakes, A. and Z. Griliches (1984): Patents and R&D at the firm level in French manufacturing: a first look. In Z. Griliches (ed.): *Research and Development, Patents and Productivity*, The University Press of Chicago, Chicago, 55–72.

[28] Parisi, M. L., Schiantarelli, F. and Sembenelli, J. (2006): Productivity, innovation and R&D: micro evidence for Italy, *European Economic Review*, 50, 2037–2061.

[29] Piekkola, H (2020): Innovation-labor biased technical change and growth – Evidence from Finnish innovation value chains. GOBALINTO (Deliverable D6.4).

[30] Power, L. (1998): The missing link: technology, investment, and productivity. *Review of Economics and Statistics*, 80, 300–313.

[31] Resutek, R.J. (2022): Is R&D Really That Special? A Fixed-Cost Explanation for the Empirical Patterns of R&D Firms. *Contemporary Accounting Research*. Vol. 39(1), 721–749.

Appendix A. Derivation of the cost-, value added-, and profit function

Consider the production function

$$Y_{it} = A^*_{it} M^{\rho}_{it} g(L_{it}, H_{it})^{\varepsilon}$$

where A^*_{it}, Y_{it} and M_{it} denote productivity, output and materials, and

$$g(L_{it}, H_{it}) = \left[a^{\frac{1}{\sigma}} L_{it}^{(\sigma-1)/\sigma} + (1-a)^{\frac{1}{\sigma}} H_{it}^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)}$$

is the CES aggregate of (man hours from) high skilled and low skilled workers, H_{it} and L_{it}. The notation in this Appendix differs in some cases from that of the main text, as we ignore quasi fixed factors (K and F) and include only two types of labour. The production function can trivially be augmented with quasi-fixed factors, e.g., K_{it}, by defining $A^*_{it} = A_{it} K^{\kappa}_{it}$.

Cost-minimization w.r.t. L_{it} and H_{it}, given factor prices for high and low skilled labor, w^l_t and w^h_t, conditional on M_{it} and Y_{it}, gives the conditional cost function

$$C(M_{it}, Y_{it}) = c_{it} \left(\frac{Y_{it}}{A^*_{it} M^{\rho}_{it}} \right)^{\frac{1}{\varepsilon}}$$

with

$$c_{it} = \left[a (w^l_{it})^r + (1-a)(w^h_{it})^r \right]^{\frac{1}{r}}$$

where $r = 1 - \sigma$. Now consider the problem of finding the cost minimizing M_{it}, given q_{Mt} and c_{it}:

$$\begin{aligned} M^*_{it} &= \arg\min_{M_{it}} q_{Mt} M_{it} + C(M_{it}, Y_{it}) \\ &= \arg\min_{M_{it}} q_{Mt} M_{it} + c_{it} \left(\frac{Y_{it}}{A^*_{it} M^{\rho}_{it}} \right)^{\frac{1}{\varepsilon}} \end{aligned}$$

The 1. order condition is

$$\ln M^*_{it} = \frac{1}{\rho + \varepsilon} (\ln Y_{it} - \ln A^*_{it}) + \frac{\varepsilon}{\rho + \varepsilon} \ln c_{it} - \frac{\varepsilon}{\rho + \varepsilon} \ln q_{Mt} + \frac{\varepsilon}{\rho + \varepsilon} \ln(\eta)$$

where $\eta = \rho/\varepsilon$. This leads to the following cost function:

$$
\begin{aligned}
C_{it}(Y_{it}) &= c_{it}\left(\frac{Y_{it}}{A^*_{it}M^{*\rho}_{it}}\right)^{\frac{1}{\varepsilon}} + q_{Mt}M^*_{it} \\
&= \left(\frac{c^{\varepsilon}_{it}Y_{it}}{A^*_{it}\left[\eta^{\frac{\rho\varepsilon}{\rho+\varepsilon}}Y_{it}^{\frac{\rho}{\rho+\varepsilon}}A^{*\frac{-\rho}{\rho+\varepsilon}}_{it}c_{it}^{\frac{\rho\varepsilon}{\rho+\varepsilon}}q_{Mt}^{\frac{-\rho\varepsilon}{\rho+\varepsilon}}\right]}\right)^{\frac{1}{\varepsilon}} + q_{Mt}\left[\eta^{\frac{\varepsilon}{\rho+\varepsilon}}Y_{it}^{\frac{1}{\rho+\varepsilon}}A^{*\frac{-1}{\rho+\varepsilon}}_{it}c_{it}^{\frac{\varepsilon}{\rho+\varepsilon}}q_{Mt}^{\frac{-\varepsilon}{\rho+\varepsilon}}\right] \\
&= q_{Mt}^{\frac{\rho}{\rho+\varepsilon}}c_{it}^{\frac{\varepsilon}{\rho+\varepsilon}}A^{*\frac{-1}{\rho+\varepsilon}}_{it}Y_{it}^{\frac{1}{\rho+\varepsilon}}[\eta^{\frac{\varepsilon}{\rho+\varepsilon}}+\eta^{\frac{-\rho}{\rho+\varepsilon}}] = \theta q_{Mt}^{\frac{\rho}{\tilde{\varepsilon}}}c_{it}^{\frac{\varepsilon}{\tilde{\varepsilon}}}A^{*\frac{-1}{\tilde{\varepsilon}}}_{it}Y_{it}^{\frac{1}{\tilde{\varepsilon}}} \qquad (22)
\end{aligned}
$$

where $\widetilde{\varepsilon} = \rho + \varepsilon$ and

$$\theta = \eta^{\frac{\varepsilon}{\rho+\varepsilon}} + \eta^{\frac{-\rho}{\rho+\varepsilon}} = \frac{\rho+\varepsilon}{\rho}\eta^{\frac{\varepsilon}{\tilde{\varepsilon}}}$$

The factor demand functions can be derived from (22) by Shepards lemma:

$$
\begin{aligned}
\ln H^*_{it} &= \frac{1}{\tilde{\varepsilon}}(\ln Y_{it} - \ln A^*_{it}) + \ln c_{it,h} \\
\ln L^*_{it} &= \frac{1}{\tilde{\varepsilon}}(\ln Y_{it} - \ln A^*_{it}) + \ln c_{it,l} \\
\ln M^*_{it} &= \frac{1}{\tilde{\varepsilon}}(\ln Y_{it} - \ln A^*_{it}) + \ln c_{it,M}
\end{aligned}
$$

and

$$
\begin{aligned}
c_{it,h} &= \theta \times q_{Mt}^{\frac{\rho}{\tilde{\varepsilon}}} \times \partial\left(c_{it}^{\frac{\varepsilon}{\tilde{\varepsilon}}}\right)/\partial w^h_t \\
c_{it,l} &= \theta \times q_{Mt}^{\frac{\rho}{\tilde{\varepsilon}}} \times \partial\left(c_{it}^{\frac{\varepsilon}{\tilde{\varepsilon}}}\right)/\partial w^l_t \\
c_{it,M} &= \theta \times c_{it}^{\frac{\varepsilon}{\tilde{\varepsilon}}} \times \partial\left(q_{Mt}^{\frac{\rho}{\tilde{\varepsilon}}}\right)/\partial q_{Mt}
\end{aligned}
$$

We see that

$$
\begin{aligned}
\partial\left(c_{it}^{\frac{\varepsilon}{\tilde{\varepsilon}}}\right)/\partial w^l_t &= \frac{\varepsilon}{\tilde{\varepsilon}}c_{it}^{\frac{-\rho}{\tilde{\varepsilon}}}\frac{\partial c_{it}}{\partial w^l_{it}} = \frac{\varepsilon}{\tilde{\varepsilon}}c_{it}^{\frac{-\rho}{\tilde{\varepsilon}}}ac_{it}^{1-r}(w^l_{it})^{r-1} \\
\partial\left(c_{it}^{\frac{\varepsilon}{\tilde{\varepsilon}}}\right)/\partial w^h_t &= \frac{\varepsilon}{\tilde{\varepsilon}}c_{it}^{\frac{-\rho}{\tilde{\varepsilon}}}\frac{\partial \ln c_{it}}{\partial w^h_{it}} = \frac{\varepsilon}{\tilde{\varepsilon}}c_{it}^{\frac{-\rho}{\tilde{\varepsilon}}}(1-a)c_{it}^{1-r}(w^h_{it})^{r-1} \\
\partial\left(q_{Mt}^{\frac{\rho}{\tilde{\varepsilon}}}\right)/\partial q_{Mt} &= \frac{\rho}{\tilde{\varepsilon}}q_{Mt}^{\frac{\rho-\tilde{\varepsilon}}{\tilde{\varepsilon}}} = \frac{\rho}{\tilde{\varepsilon}}q_{Mt}^{\frac{-\varepsilon}{\tilde{\varepsilon}}}
\end{aligned}
$$

In particular we obtain

$$\ln M^*_{it} = \frac{1}{\tilde{\varepsilon}}(\ln Y_{it} - \ln A^*_{it}) + \frac{\varepsilon}{\tilde{\varepsilon}}\ln c_{it} - \frac{\varepsilon}{\tilde{\varepsilon}}\ln q_{Mt} + \frac{\varepsilon}{\tilde{\varepsilon}}\ln(\eta) \qquad (23)$$

Each firm faces a demand function

$$Y_{it} = \Phi_{it}P_{it}^{-e} \Leftrightarrow P_{it} = \Phi_{it}^{\frac{1}{e}}Y_{it}^{\frac{-1}{e}}. \qquad (24)$$

with optimal Y_{it} given by:

$$\begin{aligned} Y_{it}^* &= \arg\max_{Y_{it}} P_{it}Y_{it} - C(Y_{it}) \\ &= \arg\max \Phi_{it}^{\frac{1}{e}} Y_{it}^{\frac{e-1}{e}} - \theta q_{Mt}^{\frac{\rho}{\widetilde{\varepsilon}}} c_{it}^{\frac{\varepsilon}{\widetilde{\varepsilon}}} A_{it}^{*\frac{-1}{\widetilde{\varepsilon}}} Y_{it}^{\frac{1}{\widetilde{\varepsilon}}} \end{aligned}$$

The first order condition is:

$$\begin{aligned} &\frac{e-1}{e} Y_{it}^{\frac{-1}{e}} \Phi_{it}^{\frac{1}{e}} = \theta q_{Mt}^{\frac{\rho}{\widetilde{\varepsilon}}} c_{it}^{\frac{\varepsilon}{\widetilde{\varepsilon}}} A_{it}^{*\frac{-1}{\widetilde{\varepsilon}}} \frac{1}{\widetilde{\varepsilon}} Y_{it}^{\frac{1-\widetilde{\varepsilon}}{\widetilde{\varepsilon}}} \\ &\Updownarrow \\ \ln Y_{it}^* = &\frac{-e\widetilde{\varepsilon}}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}} \ln(\nu) + \frac{\widetilde{\varepsilon}}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}} \ln \Phi_{it} - \frac{\rho e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}} \ln q_{Mt} \\ &- \frac{\varepsilon e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}} \ln c_{it} + \frac{e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}} \ln A_{it}^* \end{aligned} \tag{25}$$

where

$$\nu = \frac{e}{e-1}\frac{\theta}{\widetilde{\varepsilon}} = \frac{e}{\rho(e-1)}\eta^{\frac{\varepsilon}{\widetilde{\varepsilon}}}$$

We do not observe output, but sales, S_{it}. Using $S_{it} = P_{it}Y_{it}^*$, we can rewrite (25) in terms of sales:

$$\begin{aligned} \ln S_{it} &= \frac{1}{e}\ln\Phi + \frac{e-1}{e}\ln Y_{it}^* \\ &= \ln(\theta_S) + \frac{1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln\Phi_{it} + \frac{e-1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln A_{it}^* \\ &- \frac{(e-1)\varepsilon}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln c_{it} - \frac{\rho(e-1)}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln q_{Mt} \end{aligned}$$

where

$$\ln(\theta_S) = \frac{-(e-1)\widetilde{\varepsilon}}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln(\nu)$$

From (23):

$$\begin{aligned} \ln M_{it}^* &= \frac{\varepsilon}{\widetilde{\varepsilon}}\ln(\eta) + \frac{1}{\widetilde{\varepsilon}}(\ln Y_{it}^* - \ln A_{it}^*) + \frac{\varepsilon}{\widetilde{\varepsilon}}\ln c_{it} - \frac{\varepsilon}{\widetilde{\varepsilon}}\ln q_{Mt} \\ &= \ln(\theta_M) + \frac{1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln\Phi_{it} + \frac{e-1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln A_{it}^* \\ &- \frac{(e-1)\varepsilon}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln c_{it} - \left(\frac{\rho e + \varepsilon(e+\widetilde{\varepsilon}-e\widetilde{\varepsilon})}{\widetilde{\varepsilon}(e+\widetilde{\varepsilon}-e\widetilde{\varepsilon})}\right)\ln q_{Mt} \end{aligned}$$

where

$$\theta_M = \eta^{\frac{\varepsilon}{\widetilde{\varepsilon}}}\nu^{-\frac{e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}$$

we obtain:

$$\ln(q_{Mt}M_{it}^{*}) = \ln(\theta_M) + \frac{1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln\Phi_{it} + \frac{e-1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln A_{it}^{*}$$
$$-\frac{(e-1)\varepsilon}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln c_{it} - \frac{\rho(e-1)}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln q_{Mt}$$

and

$$V_{it} = A_{it}^{*\frac{e-1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\Phi_{it}^{\frac{1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}c_{it}^{-\frac{(e-1)\varepsilon}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}q_{Mt}^{-\frac{\rho(e-1)}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}(\theta_S-\theta_M)$$

Variable factor costs are given by:

$$\ln C(Y_{it}^{*}) = \ln(\theta) + \frac{\rho}{\widetilde{\varepsilon}}\ln(q_{Mt}) + \frac{\varepsilon}{\widetilde{\varepsilon}}\ln(c_{it}) - \frac{1}{\widetilde{\varepsilon}}\ln A_{it}^{*} + \frac{1}{\widetilde{\varepsilon}}\ln(Y_{it}^{*}) =$$
$$\ln(\theta_C) - \frac{(e-1)\varepsilon}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln c_{it} - \frac{\rho(e-1)}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln q_{Mt} + \frac{e-1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln A_{it}^{*}$$
$$+\frac{1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln\Phi_{it}$$

where

$$\ln(\theta_C) = \ln(\theta) - \frac{e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}\ln(\nu)$$

Profit can be written as:

$$\Pi_{it} = A_{it}^{*\frac{e-1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\Phi_{it}^{\frac{1}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}c_{it}^{-\frac{(e-1)\varepsilon}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}q_{Mit}^{-\frac{\rho(e-1)}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}(\theta_S-\theta_C)$$

and profit as a share of V_{it}:

$$\frac{\Pi_{it}}{V_{it}} = \frac{(\theta_S-\theta_C)}{(\theta_S-\theta_M)} = \frac{\nu^{\frac{-(e-1)\widetilde{\varepsilon}}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}} - \frac{\widetilde{\varepsilon}}{\rho}\eta^{\frac{\varepsilon}{\widetilde{\varepsilon}}}\nu^{\frac{-e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}}{\nu^{\frac{-(e-1)\widetilde{\varepsilon}}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}} - \eta^{\frac{\varepsilon}{\widetilde{\varepsilon}}}\nu^{\frac{-e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}} = \frac{\nu^{\frac{-e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\left(\nu - \frac{\widetilde{\varepsilon}}{\rho}\eta^{\frac{\varepsilon}{\widetilde{\varepsilon}}}\right)}{\nu^{\frac{-e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\left(\nu - \eta^{\frac{\varepsilon}{\widetilde{\varepsilon}}}\right)}$$
$$= \frac{\nu^{\frac{-e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\eta^{\frac{\varepsilon}{\widetilde{\varepsilon}}}\left(\frac{e-(e-1)\widetilde{\varepsilon}}{\rho(e-1)}\right)}{\nu^{\frac{-e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\eta^{\frac{\varepsilon}{\widetilde{\varepsilon}}}\left(\frac{e-(e-1)\rho}{\rho(e-1)}\right)} = \frac{\rho\left(\frac{e}{e-1}\right)^{\frac{-e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\varepsilon^{\frac{-\varepsilon(1-e)}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\left(\frac{e-(e-1)\widetilde{\varepsilon}}{\rho(e-1)}\right)}{\rho\left(\frac{e}{e-1}\right)^{\frac{-e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\varepsilon^{\frac{-\varepsilon(1-e)}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\left(\frac{e-(e-1)\rho}{\rho(e-1)}\right)}$$

Equation (5) follows with:

$$\widetilde{\theta} = \ln\left(\left(\frac{e}{e-1}\right)^{\frac{-e}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\varepsilon^{\frac{-\varepsilon(1-e)}{e+\widetilde{\varepsilon}-e\widetilde{\varepsilon}}}\left(\frac{e-(e-1)(\rho+\varepsilon)}{(e-1)}\right)\right) \tag{26}$$

and the profit share is:

$$\frac{\Pi_{it}}{V_{it}} = \frac{e-(e-1)(\rho+\varepsilon)}{e-(e-1)\rho} \in (\frac{1-\rho-\varepsilon}{1-\rho}, 1) \tag{27}$$

Appendix B. Data sources

Accounts statistics: All joint-stock companies in Norway are obliged to publish company accounts every year. The accounts statistics contain information obtained from the income statements and balance sheets of joint-stock companies, in particular, the information about operating revenues, operating costs and result, labor costs, the book values of a firm's tangible fixed assets at the end of a year, their depreciation, and write-downs.

The structural statistics: The term "structural statistics" is a general name for statistics of different industrial activities, such as manufacturing, building and construction, wholesale and retail trade statistics, etc. They all have the same structure and include information about production, input factors, and investments at the firm level. These structural statistics are organized according to the NACE standard and are based on General Trading Statements, which are given in an appendix to the tax return. In addition to some variables, which are common to those in the accounts statistics, the structural statistics contain data about purchases of tangible fixed assets and operational leasing. These data were matched with the data from the accounts statistics. As the firm identification number here and further we use the number given to the firm under registration in the Register of Enterprises, one of the Brønnøysund registers, which has operated from 1995.

R&D statistics: R&D statistics are the survey data collected by Statistics Norway every second year up to 2001 and annually from then on. These data comprise detailed information about firms' R&D activities, in particular, about total R&D expenses with division into internally performed R&D and externally performed R&D services, the number of employees engaged in R&D activities and the number of man-years worked in R&D. In each wave, the sample is selected with a stratified method for firms with 10–50 employees, whereas firms with more than 50 employees are all included. Strata are based on industry and firm size. Each survey contains about 5000 firms, although many of them do not provide complete information.

Register of Employers and Employees (REE): The REE contains information obtained from employers. All employers are obliged to send information to the REE about each individual employee's contract start and end, working hours, overtime and occupation. An exception is made only if a person works less than four hours per week in a given firm and/or was employed for less than six days. In addition, this register contains identification numbers for the firm and the employee, hence, the data can easily be aggregated to the firm level.

National Education Database (NED): The NED gathers all individually based statistics on education from primary to tertiary education and has been provided by Statistics Norway since 1970. We use this data set to identify the length of education. For this purpose, we utilize the first digit of the NUS variable. This variable is constructed on the basis of the Norwegian Standard Classification of Education and is a six-digit number, the leading digit of which is the code for the educational level of the person. According to the Norwegian standard classification of education (NUS89), there are nine educational levels in addition to the major group for "unspecified length of education". Education levels are given in Table 4.

Appendix C: Supplementary tables

Table 4: Educational levels

Tripartition of levels	Level	Class level
	0	Under school age
Primary education	1	1st – 7th
	2	8th – 10th
Secondary education	3	11-12th
	4	12th – 13th
	5	14th – 17th
Postsecondary education	6	14th – 18th
	7	18th – 19th
	8	20th+
	9	Unspecified

Table 5: Descriptive statistics for the main variables used in the final sample[a)]

Variable	Obs	Mean	Median	Inter quartile range	
V_{it}/L_{it}	47,739	2,045	881.5	647.7	1,272
$I_{it}^{b)}/L_{it}$	30,423	148.7	35	0	152.8
L_{it}	47,739	104.7	35	17	84
$L_{it}^{(1)}/L_{it}$	47,739	.30	.27	.13	.45
$L_{it}^{(2)}/L_{it}$	47,739	.58	.59	.46	.71
$L_{it}^{(3)}/L_{it}$	47,739	.12	.04	0	.14
$I_{i,t-1}/F_{it}^{b)}$	30,423	.18	.14	0	.18
$F_{it}^{b)}/V_{it}$	30,423	2.93	.25	.05	.97
K_{it}/V_{it}	47,739	3.08	.29	.07	.88
Π_{it}/V_{it}	47,739	.74	.26	.09	.56

$^{a)}L$ is number of man-years; $L^{(m)}$ is man-years in skill category m; V, I, F, K and Π denote value added, $R\&D$ investments, R&D capital stock, physical capital stock and profits in NOK 1000 (fixed 2017 prices), resp.

$^{b)}$Conditional on $F_{it} > 0$