Choosing Technologies: Benefits of Developing Fourth Industrial Revolution Technologies

Bettina Peters* and Markus Trunschke◊

February 2022

- Preliminary draft -

Abstract

The Fourth Industrial Revolution (4IR) presents a major technology transformation towards a data-driven economy, associated with both high costs and potential benefits. We build and estimate a dynamic discrete technology choice model to explain a firm's decisions to engage in the development of new 4IR or non-4IR technologies. The model accounts for the endogenous nature of these decisions and allows them to persistently affect the firm's future productivity path. We estimate the benefits and costs of either technology using a panel data set of high-tech manufacturing firms in Germany between 2008-2016 in combination with patent information on the type of technology. We find that firms achieve a short-run average productivity increase of 7.2 % from developing 4IR technologies, 5.1 % from non-4IR technologies, and 8.8 % from doing both. Long-run average expected benefits arising through a strongly persistent productivity process are substantially higher for 4IR (118m Euro) than for non-4IR technologies (70m Euro). However, 4IR development costs are more than double non-4IR development costs. Especially for inexperienced firms, a combination of substantially higher development costs and lower expected long-run benefits constitute a high entry barrier to starting 4IR technology development. A subsidy for 4IR technology development shifts activities from non-4IR to 4IR while increasing overall development activities.

Keywords: 4IR technology, patents, R&D, dynamic structural model, industry 4.0

JEL Classification: O33, O31, C35, D25

Acknowledgments: This paper is an outcome of the GrowInPro project which was funded by the European Commission, Research Directorate General as part of the European Union Horizon 2020 Research and Innovation Action under grant agreement No 822781. Further details can be found at: www.growinpro.eu. We are grateful to Thorsten Doherr for helping us in merging MIP and EPO patent data. Any errors remain those of the authors.

*ZEW–Leibniz Centre for European Economic Research Mannheim, MaCCI, University of Luxembourg. bettina.peters@zew.de

◊ZEW–Leibniz Centre for European Economic Research Mannheim, MaCCI, KU Leuven. markus.trunschke@zew.de

1 Introduction

The Fourth Industrial Revolution (4IR) describes the ongoing process of automation and digitization in the production of goods and services using modern smart technologies which analyze and diagnose problems via the transmission and evaluation of large amounts of data without human involvement. The ongoing automation is achieved through the integration of large-scale machine-to-machine communication, the internet of things and internet of services. This represents the fourth major technology transformation after the spread of steam and water power, the mass production related to electricity, and the automation in the course of the spread of information and communication technologies (ICT). It is currently seen as one of the most important opportunities for firms to achieve long-term benefits. However, it also poses many challenges for firms as it usually leads to significant changes in firms' production processes, organizational structures, business models, and goods and services offered. In view of the growing technological opportunities stemming from 4IR related technologies and their challenges, firms have to choose, when deciding on their R&D activities, whether to invest in the development of new 4IR related technologies, focus on the development of new non-4IR related technologies, do both or nothing at all. In order to do so, firms will compare the expected benefits that they can achieve with different types of technologies with the costs of developing these new technologies. They will decide to invest in the development of new 4IR related technologies if the expected benefits are larger than associated development costs. In this paper, we build and estimate a dynamic structural discrete choice model of technology development that rationalizes this firm behavior and provides first evidence on the benefits and costs of developing new 4IR related and non-4IR related technologies.

4IR related technologies have been increasingly developed and incorporated in firms' organizational structures and production processes. Benefits are expected to arise among others from higher flexibility in production (Bartel et al. 2007), better-informed decision making (Brynjolfsson et al. 2011), and offering better customized and personalized goods and services. These different types of benefits are likely to show up in higher productivity and in turn in revenues and profits. Despite the increasing usage of 4IR related technologies, there is only scarce evidence on their productivity impact. Prior studies have mainly studied the productivity effects of information and communication technologies (ICT), which have been the backbone of the third industrial revolution. Early studies have not found significant benefits of investments in these type of digital technologies (Loveman 1994). But more recent studies uncovered increasing performance benefits of digital technologies (Bharadwaj et al. 1999; Brynjolfsson and Hitt 2003). These studies differ in their sample of firms and estimation methods used and most studies focus on specific inventions like broadband internet which hinders a general comparison of the benefits of investing in digital and non-digital technologies. Furthermore, most of these studies look at the productivity effects of adopting new information and communication technologies. Our focus is more narrow as we specifically investigate a firm's decision to invest in own research and development of such a technology, which is associated with higher risk but potentially also with higher benefits. Finally, none of these studies provide estimates of the long-term benefits of developing 4IR related and non-4IR related technologies and their development costs.

We aim to fill this gap by building and estimating a structural model of technology development choices, explicitly modelling the firm's decision process to invest in the development of either new 4IR related or non-4IR related technology. Most prior literature has employed the knowledge production function approach by Griliches (1979) to estimate the returns to R&D. This usually means including an R&D capital stock or R&D investment as an additional input factor in the production function, in our case distinguishing between a 4IR and non-4IR capital stock. In contrast, we build on recent work by Aw et al. (2011) and Peters et al. (2017), who model the decision to invest in R&D in general in a dynamic structural model. We extend their model idea by allowing firms to decide upon developing two different types of technology: 4IR and non-4IR ones. The return for the development a new 4IR or non-4IR technology is defined in this type of model as the associated long-term benefit minus the development costs for each technology.

We model the firm's choice of investing in the development of 4IR related and non-4IR related technology as discrete choices with dynamic implications on firm performance. A key component of our model is the endogenous dynamic productivity process that is affected by the firm's decision to invest in developing new 4IR related and non-4IR related technologies. Both types of technologies allow firms to get on a different, permanently higher productivity path. This path change impacts profits in the following periods and creates additional long-run benefits for the developing firm that can pay off development costs. Explicitly modeling these dynamics allows us to derive simple technology development decision rules, that will be used in further research to analyze firm behavior in simulated counterfactual firm environments such as subsidies for R&D in specific technology fields.

We estimate the model parameters using a combination of firm-level panel data on German high-tech manufacturing firms from the Mannheim Innovation Panel (MIP) and patent application data from the European Patent Office (EPO). The MIP constitutes the German contribution to the European-wide Community Innovation Survey (CIS) and provides yearly firm-level information such as revenues, capital stocks, and expenditures on intermediary inputs necessary to estimate productivity. Information about firms' technology choice is measured using patent information. We merge information on firms' patent applications at the EPO to our MIP data and exploit a recently developed patent classification scheme that categorizes patents according to protecting the invention of a 4IR related or non-4IR related technology. Our estimation sample covers more than 3400 firm-year observations from more than 1300 firms over the period 2008 to 2016.

Our results show that on average firms benefit from developing both new 4IR and non-4IR technologies. More importantly, the average boost in productivity is of higher magnitude for 4IR than for non-4IR technologies. Developing a new 4IR technology increases firm productivity in the next period, on average, by 7.2 % compared to 5.1 % for non-4IR related technology. Furthermore, we find evidence that both types of technologies are substitutes. That is investing in both types of technologies simultaneously raises productivity only by 8.8 %. These short-run productivity improvements are carried over to future periods through a highly persistent productivity evolution process over time in which productivity gains only slowly depreciate. These two channels are important for defining the long-term benefits of developing new 4IR and non-4IR related technologies. We estimate benefits for 4IR technology development to be 119m Euro, 48m Euro higher than for non-4IR technology. These high long-run expected benefits of 4IR are mainly caused by firms with previous experience in the development of 4IR technology. These firms yield on average 288m Euro long-run benefits from developing 4IR technology while the expected benefits for firms starting to develop 4IR technology is only about one third. Not only the benefits but also the costs for the development of both technologies differ substantially. For firms without experience in developing 4IR technologies, start-up costs are on average more than double the corresponding start-up costs for non-4IR technologies. Thus, the barriers to enter are much higher for the development of new 4IR technologies than for non-4IR technologies. But once the firm has incurred these sunk start-up costs, the costs of continuing the development of both technology types drop substantially to about 10 % of startup costs. But continuation costs for 4IR technologies remain higher than for non-4IR related technologies.

These higher costs for 4IR technology development present a major obstacle for firms engagement in 4IR technology development. We therefore simulate firm's reaction to a 25 % subsidy on development costs of 4IR technology leaving non-4IR technology development costs unchanged. Firms react to this counterfactual policy change by increasing 4IR technology development activity. The share of firms with 4IR would raise by 2.5 percentage points. Compared to the actual share, this corresponds to a 29 % increase. At the same time the subsidy reduces non-4IR technology development, but the overall the share of firms with development activities would rise.

In the next section, we develop the theoretical model for explaining a firm's decisions to invest in the development of new 4IR and non-4IR technologies. Section 3 explains our empirical approach to estimate the model parameters as well as the estimation method. The data and descriptive statistics are presented in section 4. Section 5 presents the empirical results, section 6 simulates a 4IR development subsidy and section 7 gives concluding remarks.

2 Theoretical Model

The theoretical model we set up in this section expresses a firm's decision to invest in developing new technologies as a discrete choice with dynamic implications. We follow the idea of Aw et al. (2011) and Peters et al. (2017) who developed a dynamic structural model to explain a firm's decision to invest in R&D. Similar to them and previous work from Doraszelski and Jaumandreu (2013), our model does not take productivity as an exogenous process. Instead, firms can affect productivity through their behavior. Specifically, our model includes investment in the development of new technologies in the productivity evolution process, allowing it to persistently impact future periods' productivity and technology development decisions. Unlike Aw et al. (2011) and Peters et al. (2017), who model the decision to invest in R&D in general, we are mainly interested in the choice between different technologies. Specifically, we distinguish between two types of technologies: 4IR and non-4IR technologies. We model the decision process in such a way that a firm sequentially decides upon investing first in non-4IR technology development and afterwards in the development of 4IR technologies.[1]

For modeling the demand side, we follow previous work by Aw et al. (2011) and Peters et al. (2017) and begin with assuming firms to operate in a monopolistically competitive market á la Dixit and Stiglitz (1977), ruling out strategic interactions between firms but allowing for short-run profits. Assuming single product firms, consumer utility maximization leads to the following firm-specific demand function

$$q_{it} = \left(\frac{p_{it}}{P_{jt}}\right)^{\eta_j} \frac{I_{jt}}{P_{jt}} e^{\phi_{it}} = \Phi_{jt} p_{it}^{\eta_j} e^{\phi_{it}}, \tag{1}$$

where p_{it} represents firm i's price for its product q_{it} at time t, while P_{jt} is a market price index of all product varieties offered in market j. I_{jt} is a measure of market size such as the aggregated disposable income that consumers are willing to spend in market j. ϕ_{it} is an exogenous firm- and time-specific demand shifter that can be interpreted to reflect differences in the product quality or consumers' desirability over product varieties of firms active in market j. We assume that ϕ_{it} is known by the firm at the time it decides upon investing in new technologies, but it is unknown to the econometrician. η_j represents the demand elasticity which we restrict to be constant within industry j, over time, and output levels. To keep notation concise, we sum up industry variables to the aggregate Φ_{jt}.[2]

On the production side, we define the firm's short-run marginal cost C_{it}^M to be a function C of physical capital K_{it}, a vector of input market prices W_{it} for variable inputs, age A_{it} and export status E_{it}, scaled by the firm's production efficiency ψ_{it}:

$$C_{it}^M = \frac{C\left(K_{it}, W_{it}, A_{it}, E_{it}\right)}{e^{\psi_{it}}} = C\left(K_{it}, W_{it}, A_{it}, E_{it}\right) e^{-\psi_{it}}. \tag{2}$$

Efficiency ψ_{it} reflects heterogeneity across firms in the technology used or in managerial abilities that lead two firms to have different output levels and thus different marginal costs even when they use the same capital input, face the same input prices, are of the same age and export status. Similar to ϕ_{it}, we assume ψ_{it} is observed by the firm but not by the econometrician.[3] An important implication of this specification is that short-run

[1] This timing assumption simplifies the calculation of development choice probabilities. An alternative timing assumption would be that a firm makes both development decisions simultaneously leading to a multinomial model as described for the general case in Aguirregabiria and Mira (2010).

[2] We use the terms market and industry interchangeably.

[3] Even though heterogeneity of firms' marginal costs are modeled only to depend on these five variables, additional sources of cost differences could be added to the model without changing general implications.

marginal costs do not depend on the firm's output level. Firms can therefore change their production output at constant marginal costs.

Firm's profit maximization then leads to a price setting rule stating that a firm sets its price p_{it} as a constant markup over its marginal costs which are independent of its output level. The markup can be expressed in terms of the demand elasticity as $\frac{\eta_j}{1+\eta_j}$, leading to the following price equation:

$$p_{it} = \frac{\eta_j}{1+\eta_j} C_{it}^M. \tag{3}$$

Assuming supply equals demand in equation (1), marginal costs are given by equation (2) and price follows equation (3), the firm's revenues R_{it} are then straight forward given as:

$$\begin{aligned} R_{it} = p_{it} q_{it} &= \left(\frac{\eta_j}{1+\eta_j}\right) C_{it}^M \Phi_{jt} \left[\left(\frac{\eta_j}{1+\eta_j}\right) C_{it}^M\right]^{\eta_j} e^{\phi_{it}} \\ &= \left(\frac{\eta_j}{1+\eta_j}\right) \Phi_{jt} \left[C\left(\cdot\right) e^{-\psi_{it}}\right] \left[\left(\frac{\eta_j}{1+\eta_j}\right) C\left(\cdot\right) e^{-\psi_{it}}\right]^{\eta_j} e^{\phi_{it}} \\ &= \left(\frac{\eta_j}{1+\eta_j}\right)^{1+\eta_j} \Phi_{jt} C\left(\cdot\right)^{1+\eta_j} e^{-(1+\eta_j)\omega_{it}} \end{aligned} \tag{4}$$

In our model, revenue differences across firms are driven by observed differences in physical capital K_{it}, input prices W_{it}, age A_{it} and export status E_{it} through their impact on firms' marginal costs, captured by $C(.)$, differences in demand elasticity η_j across industries as well as by differences in unobserved firm-specific production efficiency ψ_{it} and demand shifter ϕ_{it}. We follow Peters et al. (2017) and summarize the unobserved components to revenue productivity ω_{it}, which is defined as $\omega_{it} = \psi_{it} - \frac{1}{1+\eta_j}\phi_{it}$. In the remainder of the paper, we call revenue productivity ω_{it} productivity for short.

Given the revenue equation (4), we can derive firm's short-run profits π_{it} as:

$$\pi_{it} = \pi(\omega_{it}) = R_{it} - C_{it}^M q_{it} = -\frac{1}{\eta_j} R_{it}. \tag{5}$$

Equation (5) implies that short-run profits are a function of all variables affecting a firm's revenues. In particular, it depends upon endogenous unobserved productivity ω_{it}.[4] In the following part of the model, we assume that firms can impact unobserved productivity ω_{it} by deciding upon investing in developing new technologies. In contrast to previous work focusing on the overall decision to invest in R&D, we are particularly interested in technology choice, that is in explaining a firm's decision to invest in developing new 4IR technologies d_{it} compared to developing new technologies in more traditional non-4IR fields n_{it}. We assume a firm sequentially decides upon investing in developing both technologies which in turn may affect its productivity development process. Importantly,

[4] For ease of representation, we omit the exogenous variables K, W, A and E and write π only as a function of the endogenous variable ω.

these investment decisions do not impact the current period's but only subsequent periods' productivity. Furthermore, we allow the productivity effects to differ across technologies. This will be an important ingredient in explaining a firm's decisions to invest in the development of the two technologies as differences in productivity effects drive differences in the expected long-run benefits of both choices. We model productivity to develop as a first-order Markov process:

$$\omega_{it+1} = g(\omega_{it}, d_{it}, n_{it}) + \xi_{it+1}. \tag{6}$$

Function $g(.)$ represents the impact of the current period's productivity level ω_{it} and technology development decisions d_{it} and n_{it} on future productivity ω_{it+1}. ξ_{it+1} is an idiosyncratic exogenous stochastic shock to productivity in period $t+1$, unknown to the firm when making its technology development decisions in period t. We furthermore assume that the productivity shocks ξ_{it+1} are *i.i.d.* across time and firms and follow a distribution f with zero mean and variance σ_ξ^2, hence: $\xi_{it+1} \sim f(0, \sigma_\xi^2)$.[5]

The level of persistence in productivity, captured by the impact of ω_{it} on ω_{it+1}, is highly important for the firm's decision to invest in productivity improvements. If persistence is high, any productivity improvement only slowly depreciates over time. This means that investments in developing new 4IR or non-4IR technologies in period t not only affect the future productivity level in period $t+1$, but are strongly carried over to subsequent periods, allowing for higher long-term returns to these investments. Low levels of persistence in turn alleviate future returns of current technology investments over a shorter time period.

The productivity evolution process described in equation (6) together with the short-run profit equation (4) allows us to derive the expected future benefits of investing in developing new 4IR and non-4IR technologies. But the firm's decision whether to invest in developing these technologies not only depends on the expected benefits gained but also on the costs associated with developing them. The decision rule for firms to invest in developing each technology type is simple: If the expected benefits exceed its costs, the firm invests in developing that technology and if benefits are lower, the firm chooses not to invest. Costs for developing 4IR technology C_{it}^d and non-4IR technology C_{it}^n can vary substantially between firms and the technology invested in. Some firms may develop a new technology at lower cost, depending on their size and previous technology development experience. The latter reflects the idea that firms differ in their capabilities to develop new technologies: Through past development experience, firms have accumulated a higher knowledge base allowing them to design future new technologies at lower costs. Therefore, firms who did not invest in developing new technology of type $\tau = d, n$ in period $t-1$ are confronted with startup costs for developing the respective technology in period t. Similarly, firms with past development experience in a specific technology τ face continuation costs for designing this type of technology in t. We model development costs C^τ of either

[5] In equation (6) the influence of past productivity is restricted to a one year lag only. One could generalize the function by allowing more distant periods to have an influence, too. However, this adds more complexity to the estimation and increases data requirements.

technology type as an $i.i.d.$ draw from an exponential distribution, allowing each mean γ^d and γ^n to vary with the firm's capital stock as a measure of firm size and with its prior technology development experience d_{it-1} or n_{it-1}:

$$\begin{aligned} C_{it}^d &\sim exp(\gamma^d(K_{it}, d_{it-1})), \\ C_{it}^n &\sim exp(\gamma^n(K_{it}, n_{it-1})). \end{aligned} \tag{7}$$

These costs are not observable to us but known to the firm when choosing to invest in either technology.

When making its technology development decisions in period t, the firm knows about its past technology decision d_{it-1} and n_{it-1} and current capital stock, age and export status. Capital, age and export status are assumed to be exogenously determined in our model. We furthermore assume that the firm observes its current productivity level ω_{it}, its short-run profit function $\pi(\cdot)$ and the productivity evolution process $g(\cdot)$. In our model, $s_{it} = (\omega_{it}, d_{it-1}, n_{it-1})$ is firm i's set of endogenous state variables. Given the exogenous variables and the current state variables s_{it}, firm i aims to maximize its value function

$$V(s_{it}) = \pi(\omega_{it}) + \delta E[V(s_{it+1}|\omega_{it}, d_{it}, n_{it})] - E[C_{it}^d]d_{it} - E[C_{it}^n]n_{it}, \tag{8}$$

consisting of the current period's short-run profits $\pi_{it}(\omega_{it})$ and the firm's expected future value $E[V(\cdot)]$, discounted by a constant factor δ, and the associated expected development costs of the firm's technology development decision. The firm's optimal value function before it observes continuation or startup costs can be expressed as

$$V(s_{it}) = \pi(\omega_{it}) + \int_{C^d}\int_{C^n} \max_{d,n\in\{0,1\}} \left\{ \begin{array}{l} \delta E_t[V(s_{it+1}|\omega_{it}, d_{it}=0, n_{it}=0)]; \\ \delta E_t[V(s_{it+1}|\omega_{it}, d_{it}=0, n_{it}=1)] - C_{it}^n; \\ \delta E_t[V(s_{it+1}|\omega_{it}, d_{it}=1, n_{it}=0)] - C_{it}^d; \\ \delta E_t[V(s_{it+1}|\omega_{it}, d_{it}=1, n_{it}=1)] - C_{it}^n - C_{it}^d \end{array} \right\} dC^n dC^d. \tag{9}$$

The expected future values at time t, $E_t[V(\cdot)]$, for all possible choices $d \in \{0,1\}$ and $n \in \{0,1\}$ are defined as

$$E[V(s_{it+1}|\omega_{it}, d_{it}, n_{it})] = \int_\omega V(s_{it+1}|\omega_{it}, d_{it}, n_{it}) dG\left(\omega_{it+1}|\omega_{it}, d_{it}, n_{it}\right). \tag{10}$$

As shown in equation (9), each period the firm chooses from those four outcome possibilities $(d,n) \in \{(0,0),(1,0),(0,1),(1,1)\}$ the option which has the highest expected future profits net of associated technology development costs. We make explicit assumptions about the timing of both development choices. Most flexible would be to allow for simultaneous decisions as described in Aguirregabiria and Mira (2010), leading to rather complicated development choice probabilities. Instead, we assume that the firm first observes its individual development costs for non-4IR technology C_{it}^n and makes its decision whether to develop a new non-4IR technology. Subsequently, the firm observe its development costs for 4IR technology and decides upon 4IR technology. This timing assumption simplifies the calculation of development choice probabilities considerably. The choice possibilities in equation (9) can be rewritten to explicitly reflect the structure of the choices during one period. After observing its non-4IR development costs the firm faces the following choice:

$$E[V(s_{it+1}|\omega_{it}, d_{it}, n_{it})] = \int_{\omega} \{V(s_{it+1}|n_{it} = 1; \omega_{it}) - C^n_{it}, V(s_{it+1}|n_{it} = 0; \omega_{it})\} \\ dG\left(\omega_{it+1}|\omega_{it}, d_{it}, n_{it}\right) \tag{11}$$

Here, the expected future value consists of the choice between two interim future value functions which depend on the firm's non-4IR technology development choice n_{it}. These conditional value functions $V(s_{it+1}|n_{it}; \omega_{it})$ themselves consist of the second choice the firm has to make in the same period, the 4IR development choice:

$$V(s_{it+1}|n_{it}; \omega_{it}) = \int_{\omega} \{E[V(s_{it+1}|d_{it} = 1, n_{it}, \omega_{it})] - C^d_{it}, E[V(s_{it+1}|d_{it} = 0, n_{it}, \omega_{it})]\} \\ dG\left(\omega_{it+1}|\omega_{it}, d_{it}, \right) \tag{12}$$

As explained earlier, differences in expected future firm value stem from the effect of the technology investment on future productivity. Expected marginal benefits of investing in developing 4IR technology can therefore be written as the difference between the expected future firm value when investing in developing 4IR technology and when not investing, conditional on the current level of state variables

$$\Delta_d E_t[V(s_{it+1})] = \delta E_t[V(s_{it+1}|\omega_{it}, n_{it}; d_{it} = 1)] - \delta E_t[V(s_{it+1}|\omega_{it}, n_{it}; d_{it} = 0)]. \tag{13}$$

Similarly, the marginal benefit of investing in developing non-4IR technology is the difference between the expected future firm values with and without developing new non-4IR technology, conditional on all firm state variables

$$\Delta_n E_t[V(s_{it+1})] = \delta E_t[V(s_{it+1}|\omega_{it}, d_{it-1}; n_{it} = 1)] - \delta E_t[V(s_{it+1}|\omega_{it}, d_{it-1}; n_{it} = 0)]. \tag{14}$$

The decision rule, that the firm will choose to develop new 4IR technology if its marginal benefit exceeds its development cost, is thus associated with the condition: $\Delta_d E_t[V s_{it+1}] > C^d_{it}$. Similarly, it will opt for developing new non-4IR technology if $\Delta_n E_t[V(s_{it+1})] > C^n_{it}$.

3 Empirical Approach

We describe in this section how we estimate parameters of the previously derived theoretical model for German firms in high technology industries. We split the estimation into two parts. The first one estimates parameters of the revenue function and the productivity evolution process. The second one uses the dynamic decision rules derived in the theoretical model to estimate latent technology development costs, choice probabilities and choice-specific expected value functions.

Revenue Function and Productivity Evolution Process

In order to obtain estimates for the parameters of the revenue function, we have to specify some additional functional forms. First, we choose the function $C(.)$ in the marginal cost equation (2) to be of the following Cobb-Douglas type form

$$C(K_{it}, W_{it}, A_{it}, E_{it}) = K_{it}^{\beta_k} W_t^{\beta_w} e^{\beta_0 + \sum_{z=1}^{Z} \beta_{a_z} A_{z,it} + \beta_e E_{it}}. \tag{15}$$

This functional form implicitly assumes that input prices W_t do not differ between firms. Variation of marginal costs with the firm's age is captured by a set of four age dummies $A_{z,it}$. We also allows for potential differences in marginal costs for exporters by including an export dummy E_{it}. Replacing marginal costs in the revenue equation (4) by the specified functional form in equation (15) and taking the natural logarithm, we end up with a basic form of the estimation equation of firm's revenue:

$$r_{it} = \lambda_{jt} + (1 + \eta_j)(\beta_0 + \beta_k k_{it} + \sum_{z=1}^{Z} \beta_{a_z} A_{z,it} + \beta_e E_{it} - \omega_{it}) + \epsilon_{it}. \tag{16}$$

Lower case letters r, k, and w denote log values of revenue R, capital K and wages W. In equation (16) all time- and industry-specific factors of the revenue function that do not vary across individuals are summarized into the expression $\lambda_{jt} = (1 + \eta_j) ln \left(\frac{\eta_j}{1+\eta_j}\right) + ln\,(\Phi_{jt}) + (1+\eta_j)\beta_w w_t$. We further include an i.i.d zero-mean error term ϵ_{it} to the revenue equation, capturing idiosyncratic transitory shocks that are unknown to the firm when it makes input and technology development decisions. In contrast, productivity ω_{it} is a variable that is unobservable to us but known by the firm at the time it makes its decisions. Estimating equation (16) without accounting for this, would lead to biased parameter estimates (Olley and Pakes 1996). Following the idea of the control function approach developed by Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015), we use material input demand to proxy for unobserved productivity. We deviate from the originally proposed approach which models material demand as a non-parametric function and define it similarly as Doraszelski and Jaumandreu (2013) and Peters et al. (2017), using the structure imposed by our theoretical model and Shephard's lemma to derive demand for material input. The log of material demand m_{it} is given by

$$m_{it} = \beta_{jt} + (1 + \eta)\beta_k k_{it} + (1 + \eta) \sum_{z=1}^{Z} \beta_{a_z} A_{z,it} + (1 + \eta)\beta_e E_{it} - (1 + \eta)\omega_{it}. \tag{17}$$

Similar to λ_{jt}, β_{jt} summarizes industry- and time-specific factors. It is defined in our model as $\beta_{jt} = ln\,(1 + \eta_j) + ln\,(\Phi_{jt}) + \eta_j\, ln \left(\frac{\eta_j}{1+\eta_j}\right) + ln\beta_w - (1 - (1 + \eta_j)\,\beta_w)\, w_t$. Inverting this equation allows us to express unobserved productivity ω_{it} as a function of observed variables

$$\omega_{it} = \left(\frac{1}{1+\eta_j}\right)\beta_{jt} + \beta_k k_{it} + \sum_{z=1}^{Z} \beta_{a_z} A_{z,it} + \beta_e E_{it} - \left(\frac{1}{1+\eta_j}\right) m_{it}. \tag{18}$$

We continue by making functional form assumptions for the productivity development process in equation (6). Defining $g(\omega_{it-1}, d_{it-1}, n_{it-1})$ to be a cubic function in the past period's productivity ω_{it-1} and linear in the technology development choices d_{it-1} and n_{it-1} leads to

$$\omega_{it} = \alpha_0 + \alpha_1\omega_{it-1} + \alpha_2\omega_{it-1}^2 + \alpha_3\omega_{it-1}^3 + \alpha_4 d_{it-1} + \alpha_5 n_{it-1} + \alpha_6 d_{it-1} n_{it-1} + \xi_{it}. \tag{19}$$

The interaction term of d_{it-1} and n_{it-1} allows for complementarities between both technology development investments. That is, a positive estimate of α_6 indicates that the firm gets an additional boost in productivity when it decides to invest in developing both new 4IR and non-4IR technology compared to developing only one technology. Whereas we would interpret a negative estimate of α_6 as evidence for both technology types being substitutes instead of complements. We then replace the lagged unobserved productivity term ω_{it-1} in (19) with observed lagged variables according to equation (18) and substitute the resulting expression for productivity ω_{it} into the revenue equation (16). This results in our final estimation equation for firm revenues

$$\begin{aligned} r_{it} = {} & \lambda_0 + \lambda_{jt} + (1+\eta_j)\beta_k k_{it} + (1+\eta_j)\textstyle\sum_{z=1}^{Z} \beta_{a_z} A_{z,it} + (1+\eta_j)\beta_e E_{it} \\ & - \quad \alpha_1 \quad [\beta_{jt-1} + (1+\eta_j)\beta_k k_{it-1} + (1+\eta_j)\textstyle\sum_{z=1}^{Z} \beta_{a_z} A_{z,it-1} + (1+\eta_j)\beta_e E_{it-1} - m_{it-1}] \\ & - \tfrac{\alpha_2}{1+\eta_j} [\beta_{jt-1} + (1+\eta_j)\beta_k k_{it-1} + (1+\eta_j)\textstyle\sum_{z=1}^{Z} \beta_{a_z} A_{z,it-1} + (1+\eta_j)\beta_e E_{it-1} - m_{it-1}]^2 \\ & - \tfrac{\alpha_3}{(1+\eta_j)^2} [\beta_{jt-1} + (1+\eta_j)\beta_k k_{it-1} + (1+\eta_j)\textstyle\sum_{z=1}^{Z} \beta_{a_z} A_{z,it-1} + (1+\eta_j)\beta_e E_{it-1} - m_{it-1}]^3 \\ & - (1+\eta_j)[\alpha_4 d_{it-1} + \alpha_5 n_{it-1} + \alpha_6 d_{it-1} n_{it-1}] - (1+\eta_j)\xi_{it} + \epsilon_{it}. \end{aligned} \tag{20}$$

In order to estimate equation (20), we first estimate the industry-specific demand elasticities η_j. According to equation (5) of our model, $\frac{C^M q}{R} = 1 + \frac{1}{\eta_j}$. We thus recover industry-specific demand elasticities by estimating the ratio of total variable cost to revenue on a constant for each industry separately, and use these estimates to back out $\hat{\eta}_j$. We then include the industry demand elasticities $\hat{\eta}_j$ as data in equation (20) and estimate the remaining parameters using nonlinear least squares (NLLS). Standard errors are estimated using heteroskedasticity- and autocorrelation robust standard errors following Davidson (2004).

Development Costs and Value Functions

In the second part of the estimation procedure we use the dynamic decision problem and its associated decision rules to estimate the remaining parameters of the model using a

Bayesian Markov chain Monte Carlo (MCMC) estimator. The basis of the estimation procedure builds the likelihood function consisting of the product of each firm's joint conditional development choice probabilities at each point in time

$$\mathcal{L}(\gamma|d_{it}, n_{it}, s_{it}) = \prod_i \prod_t P(d_{it}, n_{it}|s_{it}, \gamma). \tag{21}$$

Because we allow the cost distributions' means to differ with a firm's previous R&D experience the parameter vector of the likelihood consists of four parameters: startup and continuation costs (γ^{sn}, γ^{mn}) for non-4IR technology development and startup and continuation costs (γ^{sd}, γ^{md}) for 4IR technology development. Using the assumed development choice timing and the assumption that startup and continuation costs are *i.i.d.* draws from a known distribution the choice probabilities are conditionally independent and can be rewritten as

$$P(d_{it}, n_{it}|s_{it}) = P(d_{it}|s_{it})P(n_{it}|s_{it}). \tag{22}$$

The decision rules in the theoretical model in section (2) show that a firm develops a specific technology if its expected marginal benefits are higher than its development costs. Therefore, we can rewrite the conditional choice probabilities of both technology types. The conditional development choice probability for non-4IR technology is then

$$P(n_{it} = 1|s_{it}) = P\Big(C^n_{it}(n_{it-1}, k_{it}) \leq E[V(s_{it+1}|n_{it} = 1; \omega_{it}) - E[V(s_{it+1}|n_{it} = 0; \omega_{it})\Big), \tag{23}$$

with a state variable vector given by $s_{it} = (\omega_{it}, k_{it}, n_{it-1}, d_{it-1})$. We model the development cost the firm faces as a draw from an exponential distribution whose mean depends on the capital stock k_{it} and previous non-4IR technology development activity n_{it-1}. Therefore, the development costs in equation (23) are a function of both variables. The conditional development choice probability for 4IR technology is analogously given by

$$P(d_{it} = 1|s_{it}) = P\Big(C^d_{it}(d_{it-1}, k_{it}) \leq E[V(s_{it+1}|d_{it} = 1; n_{it}, \omega_{it}) - E[V(s_{it+1}|d_{it} = 0; n_{it}, \omega_{it})\Big). \tag{24}$$

An important difference to note is that the future period's value functions are not just conditional on the current periods 4IR development decision. They also condition on the current period's non-4IR development decision because of the timing assumptions for development choices. As stated in Aguirregabiria and Mira (2010) and Aw et al. (2011) because of the assumed exponential distributions of development costs simple closed form expressions exist for equation for both choice probabilities which we make use of in the estimation procedure.

The conditional development choice probabilities of each firm depend on the value functions (9) and expected value functions (10), (11), (12). We estimate those alongside the development cost parameters by iterating this system of equations as described in Rust (1987) and Das et al. (2007). In this approach value functions are iterated until they converge to a fixed point for each calculation of the likelihood function. Following Aw et al. (2011), we employ a Bayesian MCMC estimator approximating the posterior distribution of each parameter instead of estimating point estimates (as in e.g. Peters et al. (2017)). Instead of maximising the likelihood function it is evaluated in each iteration of the Bayesian MCMC algorithm for a given set of cost distribution parameters γ. We choose this approach instead of using a maximum likelihood estimation because it is less affected by problems arising from local maxima in likelihood functions of complex models. For each of the four cost distribution parameters we choose highly diffuse normally distributed priors with a zero mean and a standard deviation of 1000. We discretize the state space $s_{it} = (\omega_{it}, n_{it-1}, d_{it-1})$ into a grid of 50 points for ω and two for each past development activity n_{it-1}, d_{it-1} leading to a grid of 200 points. We allow the value functions to vary with a number of firm properties by additionally defining 50 firm types for capital stock, 5 for a firm's industry, 4 for firm age, and two for a firm's export status. This amounts to a total of 2,000 firm types each having their own state space. We use cubic splines to interpolate between the grid points for capital and productivity for each industry, age and export group when calculating each observation's payoff. We run the Bayesian MCMC algorithm for 17,000 iterations. After an initial burn-in period of 1,000 iterations, we are therefore left with a sample of 16,000 values from the posterior distributions of each cost parameter.

4 Data

We use two data sources to estimate our structural model. The Mannheim Innovation Panel (MIP) provides detailed information on all firm-specific variables necessary to estimate firm revenue and productivity, except for the information on technology choice. In order to identify whether a firm has invested in developing new 4IR and non-4IR technology, we therefore rely on recent patent information from the Worldwide Patent Statistical Database (PATSTAT) that allows us to classify patents into these two technology types and merge the patent data to the MIP.

The MIP is a representative survey on innovation activities in firms in Germany. It constitutes the German contribution to the European-wide harmonized Community Innovation Survey (CIS). In Germany, the innovation survey has been conducted on a yearly basis by the ZEW - Leibniz Centre for European Economic Research since 1993. The MIP follows the Oslo Manual, stating guidelines for measuring innovation activities in firms.[6] The sample is stratified by firm size (8 size classes), region (East/West Germany), and industry (56 two-/three digit Nace classes).[7] The MIP covers firms belonging to manu-

[6]See OECD and Eurostat (2019) for a comprehensive description of the Oslo Manual.

[7]The samples are drawn from the Creditreform database which is the largest credit rating agency in Germany. The Creditreform database is used because the official business register is not accessible to the public in Germany. The Creditreform database includes a stock of more than 8 million firms of which about

facturing, mining, energy and water supply, wholesale, transportation, information and communication, financial as well as other business-related services. Firms once drawn into the sample are asked to participate in the following years to create a panel structure. Sampling is updated every two years to account for firms exiting or entering the selection criteria of the target population. Survey participation is voluntary and the average response rate is about 25%, leading to an unbalanced panel data set.[8] The MIP contains firm-specific information on variables like sales, export, employment, labor and material costs, physical assets and investment, and variety of firm-specific information on innovation activities like R&D expenditure, product and process innovation and variables describing the innovation management processes within firms. More specifically, in order to estimate equation (20), we define revenue r_{it} as log of firm revenue in year t and capital k_{it} as log of the stock of fixed assets at the beginning of year t. The latter is direct information from the survey. Material cost m_{it} is measured as the log of expenditures for material, intermediate inputs, and energy in year t. To estimate the demand elasticity, we furthermore define total variable cost as the sum of material cost and labor cost. According to our model, marginal costs might also depend on the firm's export activities E_{it} and firm age. We measure export activities using an export dummy that equals 1 if the firm was an exporter in a given period t and firm age using four dummy variables $A_{1,it}$ to $A_{4,it}$ for firms being up to 9, 10-19, 20-49 and 50 and more years old.

However, despite the rich set of innovation indicators, the MIP does not allow us to identify a specific technology choice over a long time horizon, in particular whether the firm has invested in the development of new 4IR or non-4IR technologies. Firms' choices to develop new 4IR and non-4IR technologies will therefore be measured based on of patent information. We assume that a firm applying for a patent of either technology in period t, has invested in the development of the respective technology type in period $t-1$. We gather patent data from PATSTAT (autumn 2019 edition). PATSTAT provides information on applicants, inventors, priority dates, and technology classes (CPC) of more than 100 million patents from more than 100 patent offices worldwide from 1977 onwards. We identify patents protecting an invention of a new 4IR technology on the basis of a new patent cartography from the European Patent Office (Ménière et al. 2017). The cartography aims to identify patent CPC classes that belong to the ongoing digitization of industrial production processes and products (4^{th} industrial revolution). To generate the cartography, the EPO asked their patent examiners to select CPC classes relevant for 4IR. Relevant in this context means in which CPC classes they would assign 4IR inventions. A full-text search of pre-defined keywords essential for the 4^{th} industrial revolution in all patents associated with those CPC classes was then conducted to verify the reliability of the cartography. The classification was only based on the EPO's patent database. Therefore, we also restrict our analysis to EPO patents.

3,3 million are still active on the market. Comparisons reveal that it covers nearly the entire population of firms in Germany, making it the most comprehensive firm database in Germany (Bersch et al. 2014). Information on German firms in Amadeus and Orbis also stems from Creditreform data.

[8]The ZEW additionally conducts a non-response survey among non-participating firms to check and correct for possible non-response biases (for further information see Behrens et al. 2017).

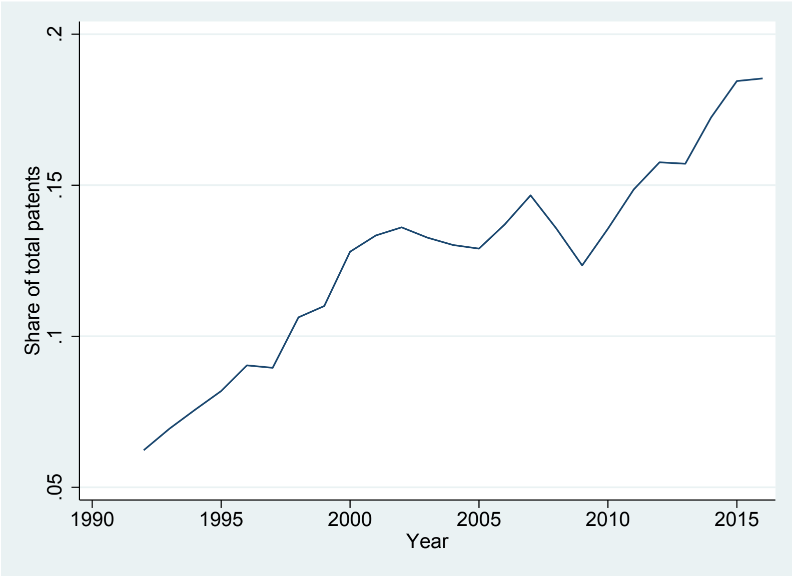
In the following, a patent is considered as a 4IR related patent if at least one of its CPC classes corresponds to one of the 320 4IR related CPC technology field ranges classified by EPO patent examiners as being relevant for 4IR inventions. These 4IR-CPC classes encompass so-called core technologies related to inventions in hardware, software, and connectivity, but also enabling technologies. Enabling technologies aim at improving the communication between the physical components and the virtual machines, and enhancing the connection between physical and data generation (e.g. GPS, visual recognition to identify patterns), transmission (interface, security), or exploitation (data analytics for big data, machine learning). Finally, 4IR related patents also contain technologies developed for final applications in various domains of the economy, including smart manufacturing, vehicles, infrastructure, home, and personal devices for individuals. If none of a patent's CPC classes fall into one of the 4IR related CPC classes, we denote it as a non-4IR related patent.

We match EPO patent information to the MIP based on firm names and addresses using a machine-learning-based search algorithm recently developed at ZEW.[9] We additionally manually check a subsample of 3.723 patents for potential inaccurate matches. For our empirical analysis, we only keep manufacturing firms observed in the MIP data and exclude firms belonging to mining, energy and water supply as well as service sectors since patenting plays only a minor role in these industries. Due to the low number of patent applications, their contribution to the identification of the technology investment parameters would be limited. In total, over the period of 1977-2018, we matched the complete history of 507,083 patents to manufacturing firms observed at least once in the MIP data, from which 235,327 patents are actually matched to firm-year observations. In the matched firm-year sample about 14.6% (34,475 patents) can be classified as 4IR related patents, and the earliest 4IR related patent application in our sample goes back to 1977, the year the EPO was founded.

Despite we match the total patent history to MIP firms, we do not consider technology choices measured by patent applications with a priority date before 1993 which is the first year of MIP data. The two most recent years of patent data show a strong decline in patent applications as shown in table A.1 in the Appendix. This decline is caused by the publication and time lag with which patents are available in PATSTAT. We therefore exclude those years from the analysis as well, so that our final panel covers the period 1993-2016. Figure 1 shows the share of 4IR related patent applications of all applications per year. At the beginning of our sample period, 4IR related patents represented a share of about 6.2 %. It has strongly increased during the second half of the 1990s. After the burst of the DotCom Bubble in 2000, the following period from 2001 to the end of the 2000s is characterized by the share of 4IR related patent applications leveling off at around the 2000 value of little more than 13 %. From 2008 onwards, which marks the new wave of digitization, the 4^{th} industrial revolution, we again see a rising number of 4IR related

[9]The *Search Engine* searches via relative relationships between search terms and candidates, while the subsequent filtering out of false positives uses absolute meta-information trained via a machine learning algorithm. This allows the search engine to yield a high recall rate while maintaining a high precision rate at the same time. The program and further information are available at https://github.com/ThorstenDoherr/searchengine.

Figure 1: Share of 4IR related patents per year



patent applications again until they reach a share of about 18.5 %.

Table 1 shows the number of patent applications and the proportion of 4IR related patents by industry. We can clearly see that the number of patents is highly unequally distributed across industries in manufacturing as is the share of 4IR related patents. Firms in the electronics sector applied for most patents overall with 61,741 patents, and they also have the highest rate of 4IR related patents among all industries with 23,5%. A very different pattern emerges for the chemical industry which has the second highest number of patent applications among all industries (53,848), but only 2.2 % (1,200 patents) are considered 4IR related patents. Machinery is ranked third in terms of patent applications (41,969), but only ranked sixth in terms of its share of 4IR related patents (7.95 %). On the contrary, the industry which has the second highest share of 4IR related patents is vehicles (22.4 %). Firms belonging to the food, textiles and mineral industry have applied for the lowest number of patents and the rate of 4IR related patents is also among the lowest in the sample. Especially the low number of 4IR related patents in these industries will limit estimates of the revenue equation at a very disaggregated industry level. We therefore restrict the estimation sample to the five high-tech sectors responsible for more than 96 % of 4IR related patent applications. The estimation procedure of the development cost parameters and the long-run benefits assumes stationarity of transition probabilities over time. However, we expect productivity benefits to change over the full sample period. Therefore, we restrict the estimation sample to the period after the financial and economic crisis 2007-2008, where most patenting activity in the area of 4IR technology can be observed. We also observe virtually no patenting activities for firms with less than 25 employees. These observations are therefore dropped from the estimation sample as well.

Table 1: Patents per industry

Industry	Non-4IR Patents	4IR Patents	% 4IR Patents Industry[a]	% 4IR Patents Total[b]	Total Patents
	Low-tech Industries				
Food	518	9	.0171	.0003	527
Textiles	1,115	88	.0732	.0028	1,203
Paper/Wood	2,565	596	.1885	.0190	3,161
Plastic	4,656	130	.0272	.0041	4,786
Minerals	2,060	69	.0324	.0022	2,129
Metal	5,158	214	.0398	.0068	5372
Misc. manuf.	887	98	.0995	.0031	985
	High-tech Industries				
Chemicals	52,648	1,200	.0223	.0382	53,848
Machinery	38,631	3,338	.0795	.1063	41,969
Electrical engineering	47,233	14,508	.2350	.4621	61,741
Instruments	12,720	3,490	.2153	.1112	16,210
Vehicles	26,526	7,656	.2240	.2439	34,182
total	194,717	31,396	0.139	1.000	226,113

Notes: [a] in % of total patents of an industry. [b] in % of total 4IR patents.
Source: ZEW - Mannheim Innovation Panel. Patstat. Own calculation.

Our theoretical model explains the firm's decision whether to invest in developing 4IR and non-4IR technologies. Assuming that a patent application is preceded by an investment in the development of that specific technology in the year before, we define d_{it} as a binary indicator that equals 1 if the firm has at least one 4IR related patent application in the following year $t+1$. Similarly, n_{it} is a dummy variable being 1 if the firm applied for at least one non-4IR related patent in year $t+1$.

Table 2 shows summary statistics for all variables used for estimating the structural model parameters. The statistics are based on our final sample that consists of 3,472 firm-year observations from 1,360 firms active in high-tech sectors. We only keep those observations with non-missing values for all variables, and we require firms to be in the sample for at least two consecutive years. The minimum number of observations per firm is 2, the maximum is 8, and on average we observe the firm 2.71 times. Concerning our main variables of interest, Table 2 shows that the share of firm-year observations with self-developed patented 4IR related technologies is 12 % while 27 % invested in developing patented non-4IR technologies.

Table 2: Summary statistics

Variable	Model	Unit	mean	med	sd	min	max
Revenues	R	mio €	539.638	51.291	1106.556	0.773	9711.470
Fixed capital	K	mio €	289.643	9.000	817.547	0.003	10284.11
Material cost	M	mio €	292.889	25.194	671.164	0.044	8508.800
Labor cost		mio €	1761.493	240	3542.372	25	41005
Total variable cost	$C^M q$	mio €	397.644	38.841	851.749	0.226	8861.522
Firm age							
0-9	A^1	0/1	0.103	0	.305	0	1
10-19	A^2	0/1	0.248	0	.432	0	1
20-49	A^3	0/1	0.370	0	.483	0	1
50+	A^4	0/1	0.278	0	.448	0	1
Exporter	E	0/1	0.944	1	0.231	0	1
Non-digital tech	n	0/1	0.276	0	0.447	0	1
Digital tech	d	0/1	0.119	0	0.324	0	1

Notes: Number of observations: 3,472. Sample period: 2008-2016. For ease of representation, all monetary variables are in million euro, for estimation we use their log values.

5 Results

We estimate and report the model parameter estimates in two parts. As described in the empirical approach in section 3, we estimate the parameters of the demand elasticities, productivity evolution process, and the cost function in the first part. These estimates are plugged in the dynamic programming problem in the second part to estimate latent development cost distribution parameters and value functions.

Revenue Function and Productivity Evolution Process

Table 3 reports the estimated demand elasticities for each industry separately. According to equations (4) and (5), the demand elasticity is an important scaling factor in how productivity translates to revenue and in turn into profits. An increase in productivity results in small profit increases for firms in industries with highly elastic demand, whereas the same productivity increase raises profits much stronger in industries with inelastic demand. All estimated demand elasticities displayed in Table 3 exhibit the expected negative sign. The estimates differ in size across industries, ranging from -3.13 in chemicals up to -4.63 in vehicles for the overall time period. Thus, we estimate that one euro revenue translates to 31.9 cents profit in chemicals on average, but only to 21.6 cents profits in vehicles.

Table 3: Demand Elasticities

Industry	Obs.	η
Chemicals	1,164	-3.13
Machinery	1,824	-4.06
Electrical engineering	1,317	-4.18
Instruments	922	-3.49
Vehicles	905	-4.63
Observations	6132	

Notes: Industry demand elasticity estimates are based on a larger sample as we only require total variables costs and revenues to be non-missing and also include firms that are observed only once or with gaps.

Using the demand elasticitiy estimates as data, we estimate equation (20) with NLLS. The results are reported in Table 4. The three polynomial coefficients of ω_{it-1} and the coefficients of d_{it-1}, n_{it-1} and their interaction term describe the productivity evolution process. We find that most coefficients are highly significantly different from zero. Only the interaction-term's coefficient is not significant at the 0.1 % level. We can draw three important conclusions from them: First, both investment in the development of new 4IR and non-4IR technologies significantly boost productivity. Developing new 4IR technology increases productivity by 7.2 % and in non-4IR technology by 5.1 %. Therefore, our model suggests that developing new 4IR technologies has a higher impact on productivity than developing new non-4IR technologies. Second, we find a negative interaction term, suggesting that both technology types are substitutes when investing into technology development. The effect of the interaction term offsets the productivity effect of the respective second technology to a non-neglible extent. That is, a firm investing in both technologies increases its productivity on average only by 8.8 %. We cannot directly compare our findings with Peters et al. (2017), because their model estimates the impact of product and process innovation, that are new to the firm, on productivity. But the evidence suggests that we find higher productivity effects of our two technology choice variables. This is reasonable as our measure of innovation implies a higher degree of novelty. Third, productivity gains are long-lived and depreciation is slow. The persistency of productivity is captured by the polynomial coefficients of ω_{t-1}. As explained above, the persistency of prior productivity levels substantially affects long-term returns to developing new 4IR and non-4IR technologies. The lower the coefficients of the ω_{it-1} terms, the shorter effects of technology development prevail in future periods. We find a non-linear relationship between the previous period's and the current productivity level because both the quadratic and the cubic terms are significantly different from zero at the 0.1 % level. The relatively high coefficients suggest that a large part of past productivity is carried over into the next period. However, the negative cubic term shows that this relationship is not monotonically increasing, but possesses decreasing marginal returns to past period's productivity.

Estimating equation (20) also provides estimates for the parameters β_k, β_{a_z} and β_e of the marginal cost function (15). Some are significantly different from zero at the 0.1 % or 1 % level. The marginal cost elasticity of capital is estimated to be $\beta_k = -0.094$. This means that an increase in the stock of capital of 10 % lowers marginal costs of production on average by 0.94 %. Based on the estimated demand elasticities, this in turn raises revenues by 2.0 %, in chemicals and 3.4 % in vehicles. The firm's export status is also negatively correlated with marginal costs. We find that an exporting firm has on average 0.01 % lower marginal costs than a non-exporter. However, this effect is not statistically significant. Our specification also allows marginal costs to vary with firm age. Compared to the reference group of young firms between 0 and 9 years, firms in the 10-19 age group have 5.1 % and firms in the 20-49 age group have 2.8 % higher marginal costs. Firms older than 49 years have 8 % lower marginal costs of production. Only the negative coefficient for the oldest age group is significantly different from zero at the 1 % level.

Table 4: Productivity Development Process and Cost Function Parameters

Variable	Coef	SE	Variable	Coef	SE
ω_{it-1}	0.403***	0.041	Capital	−0.094***	0.004
ω_{it-1}^2	0.342***	0.017	A2: 10 − 19	0.051	0.029
ω_{it-1}^3	−0.069***	0.003	A3: 20 − 49	0.028	0.027
d_{it-1}	0.072**	0.022	A4: 50+	−0.080**	0.030
n_{it-1}	0.051***	0.007	Export	−0.001	0.016
$n_{it-1} \cdot d_{it-1}$	−0.035	0.024			
			λ_0	1.382***	0.101
			Machinery	−0.015***	0.049
			Electronics	−0.020***	0.054
			Instruments	0.049***	0.030
			Vehicles	−0.005***	0.074
SE(ξ)	0.098				
Observations	3472				

Notes: Significance at the * 5% level, ** 1% level, *** 0.1% level. Time dummy variables are included in estimation but not reported. Food industry dummy is excluded as reference category.

Dynamic Parameters

The second part of the estimation procedure uses the derived likelihood to retrieve estimates of development cost distribution means. Since we use a Bayesian MCMC approach, the results represent means and standard errors of the posterior distribution. We present key parameters of the estimates in Table 5. The histogram of the posterior distributions is shown in Figure A1 in the appendix.

Table 5: Development Cost Distribution

		Mean	SD
Non-4IR technology	γ^{sn}	30.5782	2.4137
	γ^{mn}	3.3167	0.1086
4IR technology	γ^{sd}	73.2272	6.5005
	γ^{md}	7.4200	0.3447

Notes: Parameters represent moments of the posterior distributions.

The resulting cost distribution parameter estimates in Table 5 show that developing 4IR technologies is substantially more costly than non-4IR technologies. The means of posterior distributions for both the startup cost parameter and the continuation cost parameter of 4IR technologies are considerably larger than for non-4IR technologies. The posterior mean for startup development costs for 4IR technologies γ^{sd} is with about 73.23 more than double the size of the posterior mean for non-4IR technologies γ^{sn} (30.58). Also, the posterior mean of the continuation costs parameter for non-4IR technologies is with about 7.42 considerably lower than its counterpart for 4IR technologies (3.32).

These results suggest that higher entry barriers exist for firms to start developing technology that is associated with the fourth industrial revolution. Once they paid the higher startup costs, development costs for continuing 4IR technology development drop substantially, but they still stay above non-4IR technology development costs. These results also rationalize the considerably lower fraction of firms in the sample deciding to start developing 4IR technologies than non-4IR technologies.

Table 6: Expected Benefits of 4IR and non-4IR Technology Development

	Mean	Median	SD	Min	Max
$\Delta EV(n_{it}\|s_{it})$	70.648	23.460	108.508	0.478	697.547
$\Delta EV(d_{it}\|s_{it})$	118.698	27.565	233.352	0.650	1877.031
Observations	3,472				

Notes: Calculations are based on equations 13 and 14. Values are in million Euro.

Table 7: Expected Benefits for Starting and Continuation of Development

	Mean	Median	SD	Min	Max
$\Delta\ EV(n_{it}\|n_{it-1}=0, s_{it})$	55.318	16.226	96.051	0.478	637.402
$\Delta\ EV(n_{it}\|n_{it-1}=1, s_{it})$	110.761	59.093	127.407	0.775	697.547
$\Delta\ EV(d_{it}\|d_{it-1}=0, s_{it})$	95.783	24.659	187.450	0.650	1702.660
$\Delta\ EV(d_{it}\|d_{it-1}=1, s_{it})$	288.421	155.681	359.493	3.986	1877.031
Observations	3,472				

Notes: Values are in million Euro.

The structure of our model allows us to calculate the expected long-run benefits for each firm if it decides to invest in 4IR and non-4IR technologies. Following the definitions in equations (13) and (14), we calculate expected benefits given the observations state variables. Table 6 presents summary statistics for expected benefits of the firms in our sample. Overall, benefits of developing 4IR technology are higher than for developing non-4IR technology. Expected benefits are highly right skewed with medians being located well below the corresponding sample mean. The average benefit of developing 4IR related technologies lies with 118m Euro about 48m Euro above average expected benefits of non-4IR technology. A similar advantage of 4IR technologies is found for the median expected benefit (27.6 vs 23.5m Euro). Table 7 shows expected benefits conditional on the firm's previous development activity. We find that expected benefits differ substantially between first time developers and firms experienced in the development of the respective technology. Expected benefits from starting to invest in developing either 4IR or non-4IR are far lower than for continuing the development. Firms starting to develop 4IR technology expect on average 95.8m Euro in long-run benefits while firms continuing the development have 288.4m Euro expected benefits. A similar pattern is also present for non-4IR technology development. First-time developers have on average 55.3m Euro expected benefits while experienced firms have about 110.8m Euro expected benefits.

6 Policy Simulation

The results of the last section reveal that 4IR technology development exhibits high short-run productivity benefits which are transferred to subsequent periods through a highly persistent productivity development process. These productivity benefits of 4IR technology are higher than for non-4IR technology. However, also the associated technology development costs are for 4IR developers substantially higher than for non-4IR developers. These high costs present obstacles for firms to develop 4IR technology preventing them to access a higher productivity path and receive substantial long-run benefits. To examine in how far a reduction in costs lifts this obstacle and incentivises firms to engage in 4IR technology development, we construct a counterfactual environment with such reduced development costs. Namely, we model a 25 % subsidy to all costs for 4IR technology development while keeping non-4IR technology development costs constant. It is likely that firms react to this change in relative prices by increasing their 4IR technology development. However, it is unclear how firms will adjust their non-4IR development activity. Non-4IR development might either fall or stay constant.

We simulate firm behaviour for environments with and without the policy for 10 years. As starting observations we choose the first observation of each firm in the sample. This exercise is repeated 50 times.

Table 8: Reduction of 4IR Development Cost

	Year 2	Year 5	Year 7	Year 10
Δ 4IR rate	0.0157	0.0249	0.0255	0.0237
Δ non-4IR rate	−0.0143	−0.0185	−0.0198	−0.0205
Δ productivity	−0.0061	−0.0059	−0.0041	0.0011

Notes: Numbers represent average differences over 50 simulations with and without the policy change.

Average differences in firm behaviour for different years are reported in Table 6. Firms react to the subsidy by increasing 4IR technology development strongly. After two years the share of firms engaging in 4IR technology development is 1.57 % higher in the counterfactual environment. This share increases of the simulated years up to 2.55 % at year 7 before it slightly decreases to 2.37 % at year 10. Non-4IR development activity decreases in response to the subsidy. After two years the share of firms developing non-4IR technology decreases by 1.43 %. This share drops constantly over time until year 10 where 2.05 % less firms develop non-4IR technology in the counterfactual environment. Even though overall R&D activity increases with the subsidy, average productivity is slightly lower in the counterfactual environment for most years. Only in the last year average productivity is 0.1 % higher with the subsidy.

7 Conclusion

The Fourth Industrial Revolution is one of the most important opportunities for achieving long-term benefits, but also a challenge for firms. It describes the ongoing process of automation and digitization in the production of goods and services using modern smart technologies. Given the increasing importance of 4IR technologies, firms have to choose when deciding on their R&D portfolio whether to invest in the development of new 4IR technologies or to stick to the development of new non-4IR technologies or to do both. In this paper, we construct and estimate a dynamic structural model explaining the firm's technology development choices. Building on recent work by Aw et al. (2011) and Peters et al. (2017), who model the decision to invest in R&D in general, our model in contrast focuses on the choice between developing two different types of technologies: 4IR and non-4IR ones.

Most prior literature includes investment in the development of new technology into a production function framework without describing the endogenous productivity process, assuming productivity to be unaffected by the decision to invest in the development of new technology. We take a different approach by explicitly modeling the firm's decision to invest in developing both 4IR and non-4IR technology separately as an endogenous decision that allows the firm to access a continuously higher future productivity trajectory. Payoffs might be realized in future periods through productivity increases, which translate into higher revenues and in turn profits. We provide a simple decision rule for a firm's technology development choice that enables us to estimate latent technology development cost distribution parameters. They will be used in further research to analyze how firms' technology development decisions are affected by changes in the economic environment.

We estimate the model parameters using a combination of firm-level panel data from the German Community Innovation Survey from 2008 until 2016 and EPO patent data, exploiting a recently developed patent classification scheme that categorizes patents into protecting the invention of a 4IR or non-4IR technology.

Our estimates show that over the whole period, investing in the development of both types of technology significantly increases firm's productivity. More importantly, the average boost in productivity for 4IR technologies is higher than for non-4IR ones. Developing a new 4IR technology increases firm productivity in the following period on average by 7.2 % compared to 5.1 % in the case of non-4IR technology. If firms decide to develop both types of technology at the same time, productivity increases by 8.8 %. This indicates that the development of 4IR and non-4IR technologies are substitutes at the firm level. These short-run productivity improvements are carried forward to future periods through a highly persistent productivity evolution process over time in which productivity improvements only slowly depreciate. These two channels are important for defining the long-term benefits of developing new 4IR and non-4IR technologies. We find that these long-run benefits for 4IR technology development are on average with 119m Euro about 48m Euro higher than for non-4IR technology. Furthermore, expected long-run benefits are substantially larger for firms experienced in the respective technology development. 4IR technology benefits for starters are with 95m Euro more than 190m Euro lower than for experienced firms. Also expected benefits of non-4IR technology development for starters are only half the size than for experienced firms (55m vs 111m).

Estimated development startup costs for 4IR technologies are more than double the size of non-4IR technologies. This suggests that substantially higher entry barriers exist for firms considering to develop 4IR technologies. Even though development costs for continuing technology development of both technology types drop strongly, 4IR technology development costs stay above non-4IR technology development costs.

Simulating a counteractual environment with a reduction of 4IR development costs by 25 % we find that firms react to the subsidy with an overall increase in development activity. Although the share of firms developing non-4IR technology constantly decreases over years, the cost reduction incentivises firms to engage in developing more 4IR technology, leading to a net increase. However, this increased development behaviour does not seem to lead to substantially higher average productivity in the sample.

Overall, our results emphasize that firms benefit from both 4IR and non-4IR technology through increased future firm productivity. While entry into the development of 4IR technologies is rewarded with higher benefits, 4IR technologies also exhibit much higher development costs for firms starting to develop these technologies compared to non-4IR technologies. In turn, firms already developing 4IR technologies and firms already developing non-4IR technologies both experience much lower costs for continuing the development of either type of technologies.

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015): "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83, 2411–2451.

Aguirregabiria, V. and P. Mira (2010): "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, 156, 38–67, structural Models of Optimization Behavior in Labor, Aging, and Health.

Aw, B. Y., M. J. Roberts, and D. Y. Xu (2011): "R&D investment, exporting, and productivity dynamics," *American Economic Review*, 101, 1312–1344.

Bartel, A., C. Ichniowski, and K. Shaw (2007): "How does information technology affect productivity? Plant-level comparisons of product innovation, process improvement, and worker skills," *The quarterly journal of Economics*, 122, 1721–1758.

Behrens, V., M. Berger, M. Hud, P. Hünermund, Y. Iferd, B. Peters, C. Rammer, and T. Schubert (2017): "Innovation activities of firms in Germany - results of the German CIS 2012 and 2014 : background report on the surveys of the Mannheim Innovation Panel conducted in the years 2013 to 2016," .

Bersch, J., S. Gottschalk, B. Müller, and M. Niefert (2014): "The Mannheim Enterprise Panel (MUP) and firm statistics for Germany," .

Bharadwaj, A. S., S. G. Bharadwaj, and B. R. Konsynski (1999): "Information technology effects on firm performance as measured by Tobin's q," *Management science*, 45, 1008–1024.

Brynjolfsson, E. and L. M. Hitt (2003): "Computing productivity: Firm-level evidence," *Review of economics and statistics*, 85, 793–808.

Brynjolfsson, E., L. M. Hitt, and H. H. Kim (2011): "Strength in numbers: How does data-driven decisionmaking affect firm performance?" *Available at SSRN 1819486.*

Das, S., M. J. Roberts, and J. R. Tybout (2007): "Market entry costs, producer heterogeneity, and export dynamics," *Econometrica*, 75, 837–873.

Davidson, R. (2004): *Econometric theory and methods*, New York, NY: OUP Oxford.

Dixit, A. K. and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *The American Economic Review*, 67, 297–308.

Doraszelski, U. and J. Jaumandreu (2013): "R & D and productivity: Estimating endogenous productivity," *Review of Economic Studies*, 80, 1338–1383.

Griliches, Z. (1979): "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, 10, 92–116.

Levinsohn, J. and A. Petrin (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70, 317–341.

LOVEMAN, G. W. (1994): "An assessment of the productivity impact of information technologies," *Information technology and the corporation of the 1990s: Research studies*, 84, 110.

MÉNIÈRE, Y., I. RUDYK, AND J. VALDÉS (2017): "Patents and the Fourth Industrial Revolution: the inventions behind digital transformation," Tech. rep.

OECD AND EUROSTAT (2019): *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation*, The Measurement of Scientific, Technological and Innovation Activities, OECD, 4 ed.

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

PETERS, B., M. J. ROBERTS, V. A. VUONG, AND H. FRYGES (2017): "Estimating dynamic R&D choice: an analysis of costs and long-run benefits," *RAND Journal of Economics*, 48, 409–437.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

A Appendix

Table A.1: Patents per year

Year	Non-4IR Patents	4IR Patents	Total Patents
1993	2,968	283	3,251
1994	4,120	384	4,504
1995	4,162	385	4,547
1996	3,492	318	3,810
1997	5,346	480	5,826
1998	6,371	911	7,282
1999	6,567	1,008	7,575
2000	7,842	1,367	9,209
2001	8,386	1,452	9,838
2002	8,709	1,555	10,264
2003	9,294	1,413	10,707
2004	9,519	1,546	11,065
2005	9,446	1,428	10,874
2006	9,828	1,489	11,317
2007	10,641	2,027	12,668
2008	9,533	1,520	11,053
2009	10,042	1,494	11,536
2010	9,423	1,468	10,891
2011	9,179	1,478	10,657
2012	9,871	1,844	11,715
2013	9,370	1,945	11,315
2014	9,240	2,107	11,347
2015	8,332	2,172	10,504
2016	8,133	1,976	10,109
2017	5,059	1,521	6,580
2018	2,087	629	2,716
total	200,852	34,475	235,327

Figure A1: Histograms of the Posterior Densities of Development Cost Distribution Means