Uddannelses- og Forskningsudvalget 2015-16
UFU Alm.del endeligt svar på spørgsmål 168
Offentligt

Analysis of the Industrial PhD Programme

Danish Agency for Science
Technology and Innovation
Ministry of Science
Technology and Innovation

>

Analysis of the Industrial PhD Programme

Published by :

The Danish Agency for Science, Technology and Innovation

Bredgade 40

1260 København K

Tlf: 35446200

Fax: 35446201

The report is produced by Iris Group for the Danish Agency for Science, Technology and Innovation

Text: CEBR – Centre for Economic and Business Research - Johan M. Kuhn, Ph.D.

ISBN: 978-87-92372-65-9 (web)

Design: Formidabel Aps

CONTENTS

EXECUTIVE SUMMARY

This report has been prepared by the Centre for Economic and Business Research (CEBR). It presents an analysis of the economic impact of the Danish Industrial PhD Programme on participating companies and on wage and career characteristics of Industrial PhD graduates.

The Industrial PhD Programme is funded by the Danish Council for Technology and Innovation and is administered by the Danish Agency for Science, Technology and Innovation (DASTI). The programme subsidises PhD studies where the student is employed in a private sector company and simultaneously enrolled as a PhD student at a university.

The analysis follows approx. 430 individuals and approx. 270 companies that have participated in the programme and for whom relevant data is available in the selected registers.

On the individual level, we compare wage income and occupation of Industrial PhDs with regular PhDs and other university level graduates.

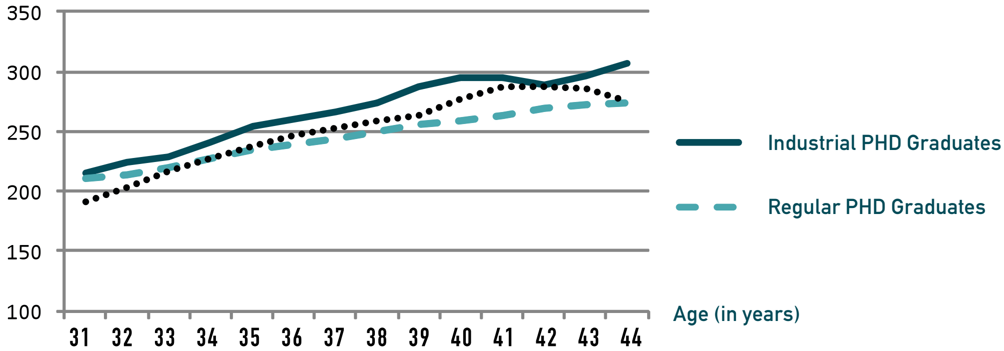
On the company level, we analyse company level developments within four success parameters:

- the number of patents applications,
- gross profit growth,
- total factor productivity, and
- employment growth.

For a sample of companies which have hosted a maximum of three Industrial PhD projects, we identify a control group of highly similar companies which have not hosted any Industrial PhD projects. We then compare developments in the success parameters in these two groups. Under identifying assumptions, the difference between the sample group and the control group isolate the causal impact of the programme on companies hosting Industrial PhD projects.

The results of the analysis can be summarised as follows: Industrial PhDs earn approx. 7-10 percent higher wages than both regular PhDs and comparable university graduates. This comparison is illustrated in FIGURE 1.

FIGURE 1: Hourly wage (DKK) in 2006, by Individual age

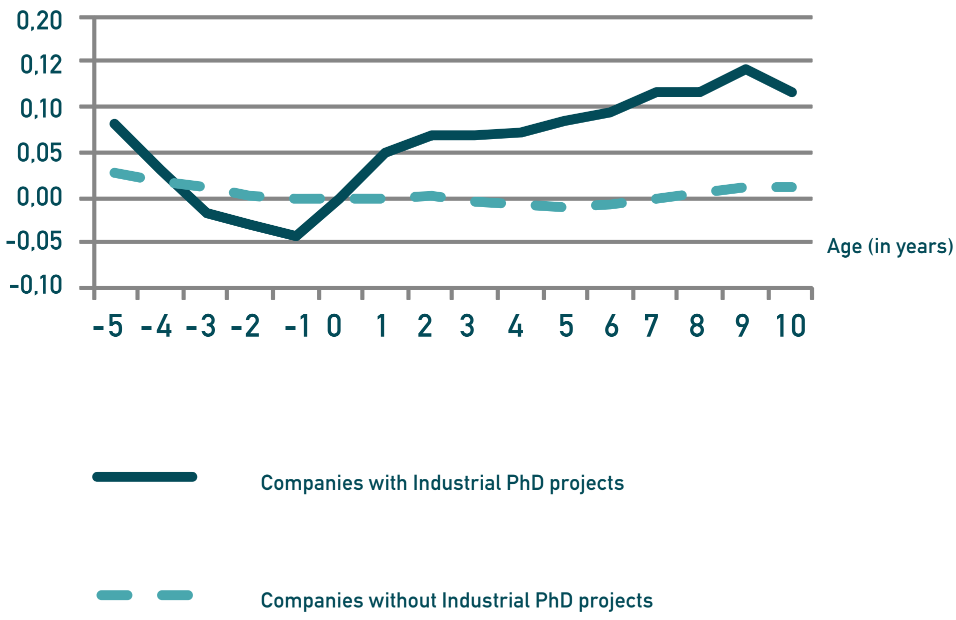


They are also more likely to be found at the top levels of their organisations' hierarchies compared to normal PhDs and more likely to be found in positions requiring high-level specialist knowledge than regular university graduates.

Companies that host Industrial PhDs see on average increasing patenting activity, illustrated by FIGURE 2. They are characterised by high growth in gross profit, and more positive developments in gross profit and employment growth than companies in the control group. We are not able to identify robust relationships between hosting Industrial PhD projects and total factor productivity developments.

FIGURE 2: Number of patent applications, high-quality matches.

Average number of patent applications per company, change relative to year before first initiating an Industrial PhD project



SAMMENFATNING (DANISH SUMMARY) >

Denne rapport er skrevet af Centre for Economic and Business Research (CEBR). Den beskriver en analyse af ErhvervsPhD-ordningens potentielle effekter på udviklingen i de deltagende virksomheder og løn- og karrieremønstre for personer, som har erhvervet deres ph.d.-grad gennem ordningen.

Et ErhvervsPhD-projekt er et treårigt erhvervsrettet ph.d.-projekt, hvor den studerende ansættes i en privat virksomhed og samtidig indskrives på et universitet.

Ved hjælp af registerdata følger analysen ca. 430 individer og 270 virksomheder, som har deltaget i ordningen. På individniveau studeres væksten i ErhvervsPhD'ernes lønindkomst i forhold til almindelige ph.d.'ere og sammenlignelige kandidater.

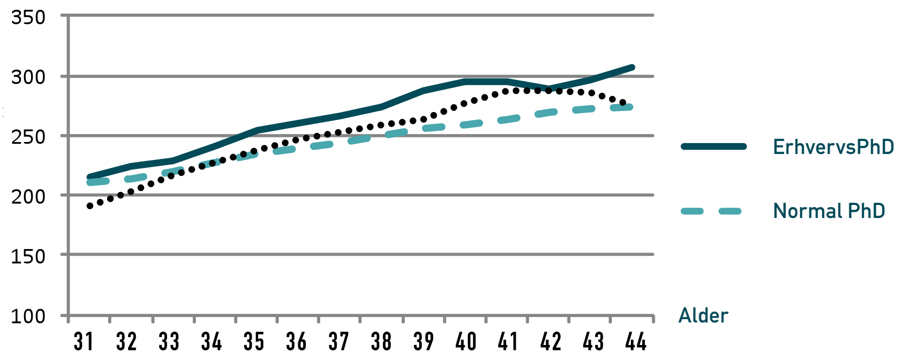
For virksomheder studeres udviklingen i patentering, bruttofortjeneste, totalfaktorproduktivitet og beskæftigelse. Hertil identificerer vi en gruppe af kontrolvirksomheder, som ikke ansætter en ErhvervsPhD, men som ellers ligner de ansættende virksomheder i størrelse, branche, alder og region.

Dermed kan vi besvare spørgsmålet om, hvorvidt de virksomheder, som ansatte en ErhvervsPhD, har haft en mere positiv udvikling i succesparametrene, end man ville have forventet på basis af udviklingen for kontrolvirksomhederne.

Analysens resultater kan sammenfattes som følger:

Efter uddannelsens afslutning har ErhvervsPhD'er i gennemsnit mellem 7 og 10 procent højere lønindkomst end normale ph.d.'er og personer med en afsluttet universitetsuddannelse. Dette er illustreret i FIGUR 1.

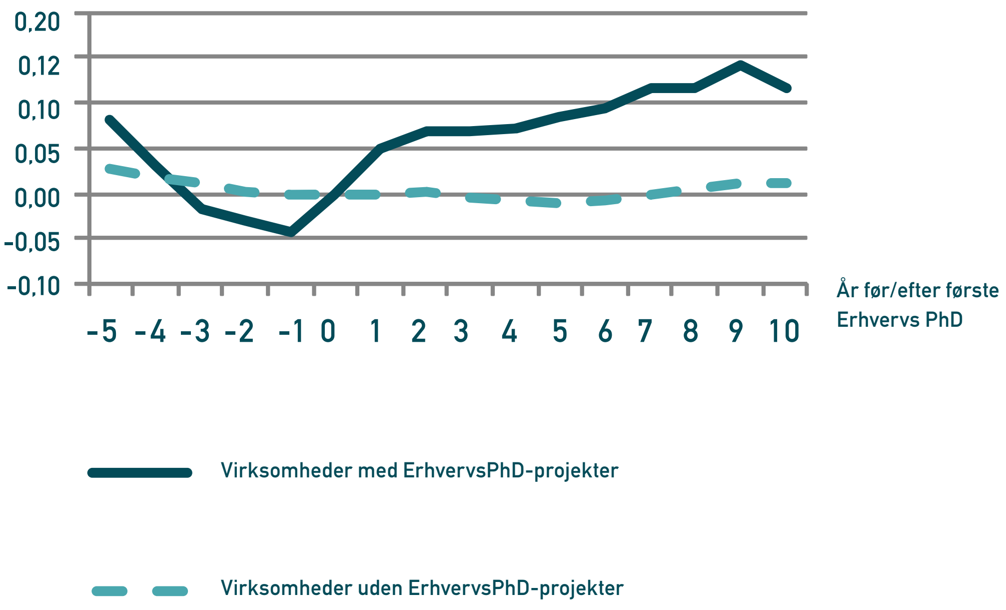
FIGURE 1: Timeløn (i kr.) in 2006, efter alder



ErhvervsPhD'ere har endvidere en væsentligt højere sandsynlighed for at blive ansat i lederstillinger end almindelige ph.d.'ere og er stærkere repræsenteret i gruppen af medarbejdere med jobfunktioner, som kræver specialviden på højeste niveau. Virksomheder, som ansætter ErhvervsPhD'ere, har i gennemsnit højere patenteringsaktivitet efter ansættelsen end før. Dette er illustreret i FIGUR 2.

FIGURE 2: Antal patentansøgniner, højkvalitetssammenligning

Gennemsnitlig antal patentansøgninger pr. virksomhed
(i afvigelser ift. året før første ErhvervsPhD-projekt)



De er også kendetegnet ved højere vækst i bruttofortjenesten/værdiskabelsen og har en mere positiv udvikling i væksten i bruttofortjenesten og medarbejderantallet end virksomhederne i kontrolgruppen.

Det er på nuværende tidspunkt ikke muligt at påvise, at ErhvervsPhD-ordningen bidrager til højere vækst i virksomhedernes totalfaktorproduktivitet.

1 INTRODUCTION

This report has been prepared by the Centre for Economic and Business Research (CEBR). It presents an analysis of the economic impact of companies participating in the Danish Industrial PhD Programme in terms of growth and value creation, and on wage income and career patterns of Industrial PhD graduates.

Even though this analysis is an evaluation of a specific Industrial PhD subsidy programme, its results might be of general interest, as programmes similar to the Danish Industrial PhD Programme have been implemented or are considered for implementation in a number of countries. However, general knowledge of their effects which can be integrated into cost-benefit analyses of these programmes is still rare.

The Industrial PhD Programme aims at increasing knowledge sharing between universities and private sector companies, promoting research with commercial perspectives, and taking advantage of competences and research facilities in private business to increase the number of PhDs.

For this purpose, the Industrial PhD students typically spend 50 percent of their time in a company and 50 percent of their time at a university while taking the degree. The Danish Agency for Science, Technology and Innovation (DASTI) subsidises the Industrial PhD's salary with a fixed monthly amount, roughly corresponding to 30-50 percent of the Industrial PhD's total salary.

The Industrial PhD programme was initiated in 1971 under the name "The Industrial Researcher Programme". In 1988 it was made possible to qualify for a PhD degree when graduating. The programme was subsequently reformed to comply with Danish PhD regulations, making every graduate a formal PhD graduate. Until 2009, approx. 1,200 projects have been started. As part of its evaluation policy, DASTI has asked CEBR to analyse the company and individual level effects of the Industrial PhD Programme. The main questions of the evaluation are whether and how participating in the Industrial PhD Programme is associated with company performance and, with regard to individuals, to what extent an Industrial PhD degree is associated with future career developments, measured by wage income and occupation.

To answer the questions outlined above, this analysis considers 430 individuals and approx. 270 companies that have participated in the programme using a matched employer-employer register dataset.

On the individual level, we compare wage income developments of Industrial PhD graduates with regular PhD graduates and individuals with a university level degree (and who have graduated at approximately the same time as the Industrial PhD graduates).

On the company level, we analyse company level developments within four success parameters:

- number of patent applications,
- gross profit growth,
- total factor productivity (TFP), and
- employment growth

Gross profit is defined as annual net sales subtracted annual costs of variable inputs (raw materials, energy, intermediate goods purchases, etc.), except labour costs. Thus, gross profit is a measure of the company's value creation.[1]

Total factor productivity is gross profit corrected for the company's use of capital and the number of employees. It is measured as the percentage-wise deviation of a company's gross profit from the gross profit that would have been expected on basis of the company's number of employees and its capital stock.[2]

To identify innovation, growth and productivity effects of hosting an Industrial PhD, we analyse increases in the number of patent applications, gross profit growth, total factor productivity and employment growth for a sample of companies which have participated in the Industrial PhD Programme. By using a control group of highly similar companies which have not participated in the programme, we can compare the developments of the success parameters of the two groups of companies to each other.

1 Gross profit is the most precise measure of the company's value creation, but one should, of course, keep in mind that a part of the company's total value creation may be passed on to consumers, may be retained in the company and increase its value (for which there is no data available for this analysis), or may take the form of positive externalities, such as knowledge and/or innovations which benefit other companies or society in general.

2 For this analysis, we measure TFP as the residuals of a Cobb-Douglas-production function estimation with total assets and the number of employees as right hand side variables.

2 DESCRIPTION OF THE INDUSTRIAL PHD PROGRAMME >

An Industrial PhD project is a three-year industrially focused PhD project where the student is hired by a company and enrolled in a university at the same time.[3]

The company receives a monthly wage subsidy of (currently) DKK 14,500 (approx. €2,000) while the university has its expenses for supervising etc. covered. The PhD student works full time on the project and divides his or her time equally between the company and the university. There are additional subsidies available for project-relevant stays abroad.

Currently, there are allocated annually approx. DKK 100-150 million (€15-20 million) for new projects. Approval rates for applications are currently above 60 percent.

Different aspects of the Danish Industrial PhD programme were addressed in earlier evaluations. DASTI (2007a)[4] concludes that Industrial PhDs are characterised by earning higher wages and are more likely to be a part of their organisation's management compared to regular PhDs. Companies hosting Industrial PhD projects expect increased patenting activity and growth.

DASTI (2007b)[5] lists several positive benefits for the participating companies. Among other things, companies may gain new knowledge, patents and licenses, growth and new market opportunities, and an increased network inside the academic world.

A similar conclusion is reached in a report by Right, Kjaer and Kjerulf from 2003. Based on interviews with participating candidates and companies in 2002, they find that a majority of companies expect the Industrial PhD to contribute to patents, while close to half of all companies expect increased earnings.

International evaluations include a report from the European University Association,[6] which concludes that participating candidates enjoy better employment opportunities due to improved skills. Two studies for the Swedish agency KK-stiftelsen[7] have also been carried out. These conclude (a) that certain conditions need to be met for projects to be successful, and (b) that the different stakeholders of Industrial PhD projects report that the programme is achieving its goals.

3 This section draws extensively on the information published by DASTI.

4DASTI, 2007a: "ErhvervsPhD - Et effektivt redskab for innovation og videnspredning".

5DASTI, 2007b: "ErhvervsPhD - Ny viden til erhvervslivet og universiteterne".

6European University Association, 2009: "Collaborative Doctoral Education - University-Industry partnerships for enhancing knowledge exchange".

7(a) KK-stiftelsen, 2003: "KK-stiftelsens företagsforskarskolor - utvärdering av ett koncept för ökat samarbete mellan akademi och näringsliv".

(b) KK-stiftelsen, 2006: "Småföretags- och institutsdoktorander för kunskaps- och kompetensutveckling".

3 INDIVIDUAL LEVEL ANALYSIS >

For the individual level analysis, information gathered by DASTI on participating individuals was merged with public register information typically referred to as the “Integrated Database of Labour Market Research (IDA)”. These data cover the period from 1980 onward and contain information on a multitude of individual demographic background characteristics, like education, gender and age.

The IDA data have – for the period 1997 to 2006 – been merged with information from the ”Wage Statistics Database”, which includes detailed information on wages and occupation, including hierarchical levels.

Also, information from education-related registers has been added to the data, to make it possible to control for inherent human capital endowments – pproximated by the grades of secondary education certificates – in the regressions.

The following analysis compares wages and careers of Industrial PhD graduates with:
(a) individuals with a university degree, but no PhD degree, and
(b) regular PhD graduates.

The validity of these comparisons depends on how similar the two groups are with the Industrial PhD graduates and the potential to control for observable factors presumably related to educational choices and, later, income and career developments. Both objectives raise some issues regarding the optimal sampling strategy, which is presented in the following:

When selecting the sample for the analysis, we obviously include all individuals who have completed an Industrial PhD education. Individuals who have completed a regular PhD education form the first control group.

With regard to university graduates, who form the second control group of individuals for comparison, there is an issue which needs to be resolved: There is a large number of secondary educations where it is not entirely clear whether they should be defined as university level educations or not.

We choose to address this issue by identifying the highest educational degrees of the Industrial PhD graduates before obtaining their Industrial PhD degrees. As a first step in the sampling procedure, we only select individuals with the same set of educations for the control group.

But without further conditions on sampling, the educational fields of the Industrial PhDs and the university graduates would be very different. For example, there would be a large share of individuals with university degrees in arts and Humanities in the control group, while these degrees are relatively uncommon in the group of Industrial PhDs. This would bias any comparison between the two groups.

For this reason, we also align the composition of the educational fields of Industrial PhDs prior to obtaining the Industrial PhD degree and the educational fields of the control group by selecting a fixed number of individuals into the control group for each Industrial PhD graduate.

Specifically, we select ten individuals into the control group for each Industrial PhD. The number ten is a compromise between being able to find individuals with the same educational degrees and a sample size large enough to isolate relationships in the data.

These individuals, referred to as the group of 'university graduates' are randomly selected, but must correspond to the educational field of the given Industrial PhD graduate (before he/she obtains her PhD degree). In the selection process, we also prefer persons of the same gender and origin (Danish vs. non-Danish), and persons who are of similar age. This way, we base the comparisons on groups of individuals similar not only in terms of their educational field, but also age, gender and origin.

The individual level analysis is based primarily on information for the year 2006, which is the last year where the data provides detailed information on wages and occupation.

3.1 Results of the individual level analysis

At present there are approx. 1,200 individuals who have participated or are participating in the Industrial PhD Programme. In year 2006, which is the last year for which all relevant data is available, 999 Industrial PhDs can be identified in the register data.

Of these, the register data shows 442 completed their projects, i.e. obtained their Industrial PhD degree, before 2006.

Additionally, there is wage information for 430 of these 442 individuals.

The wage concept used in the following analysis is Statistics Denmark's 'nw'-variable of the Wage Statistics Database. This variable is a description of the person's hourly wage income excluding pension contributions and cleaned for peculiarities such as overtime, dirty work premiums, etc.

Career developments are measured by Statistics Denmark's 'disco'-variable, also from the Wage Statistics Database. This variable categorises occupation by different hierarchical levels and work functions. The question to be considered is whether Industrial PhDs are over- or underrepresented in leadership positions (disco code 1000-1999) or positions which require high-skilled specialist knowledge – these will be denoted as specialist positions in the following (disco code 2000-2999).

Descriptive statistics

In this subsection, we describe the gross sample of all individuals associated with the Industrial PhD Programme – with or without completed Industrial PhD degrees - and of all individuals with a PhD degree in the last year in which they are observed, and all individuals selected for the group of university graduates. This ensures the most comprehensive description of these groups. However, when we turn to the comparison of wages and career patterns, we will concentrate on those individuals with completed educations (either university or PhD) in 2006.

When taking a first look at the data (TABLE 3.2.1), we find that in 2006, the last year for which data is available for this analysis, Industrial PhDs earn approx. 10 percent lower wages than regular PhDs, but are, in the current sample, almost eight years younger on average.

Industrial PhDs have slightly lower grades than normal PhDs in their secondary education examinations, but this difference is negligible relative to this variable's variation.

About four percent of the Industrial PhDs are represented at the top of their organisation's hierarchy, which is very similar to the two control groups. Approx. 60 percent of Industrial PhDs work in specialist positions, a share which locates them between regular PhDs and university graduates, where this is the case for 74 and 44 percent, respectively. Obviously, we can expect differences in both wages and positions to increase when focusing only on individuals with completed educations in the next subsection.

TABLE 3.2.1: Descriptive statistics of the individual level data (2006), mean values

	Industrial PhD students and graduates	Regular PhD students and graduates	University graduates	All
Hourly wage (DKK)	228,61	223,44	223,44	243,06
Female	0,35	0,34	0,35	0,34
Age (year)	34,55	42,66	34,72	38,98
Grade of university-entrance diploma (standard deviation: 8.9)	91,83	92,63	87,33	90,07
Non-Danish origin	0,06	0,08	0,05	0,07
Leadership position	0,04	0,04	0,03	0,04
Specialist position	0,58	0,74	0,44	0,61
Number of observations	999	12369	9625	22993

Cf. TABLE 3.2.2, we find that approx. 38 percent of all Industrial PhDs had a degree in engineering, and another approx. 23 percent had degrees in chemistry or electronics engineering before receiving their Industrial PhD degree. [8]

[8] Obviously, the group of university graduates is supposed to only consist of individuals who actually have graduated. Thus, when we formally compare the different groups of individuals in the results subsection, please note that we will not consider individuals registered as having a university-entrance diploma as their highest educational degree in 2006.

TABLE 3.2.2: Highest educational degree in 2006 (for Industrial PhD and regular PhD graduates: highest degree before receiving the PhD degree), in percent

	Industrial PhDs	Regular PhDs	University Graduates	All
Master's in engineering	37,99	12,38	32,95	20,99
Unknown	3,89	30,24	0,08	18,08
Master's in medical science	2,52	25,42	4,01	16,61
Master's in biology	9,84	14,26	11,16	12,94
Master's in chemical engineering	10,53	5,3	11,42	7,76
Master's in electronics engineering	12,81	3,46	12,02	6,99
University-entrance diploma	9,38	0,6	15,79	6,54
Master's in pharmaceutics	6,41	4,04	6,73	5,13
Master's in biochemistry	6,64	4,3	5,84	4,96

Before turning to wage and career comparisons, we take a brief look at the kind of PhD degrees of Industrial and regular PhDs - see TABLE 3.2.3 for the most popular subjects. We find these two groups to be quite different in their compositions of the specific degrees. Consequentially, later comparisons of wages and careers between Industrial and regular PhDs will have to take these differences into account.

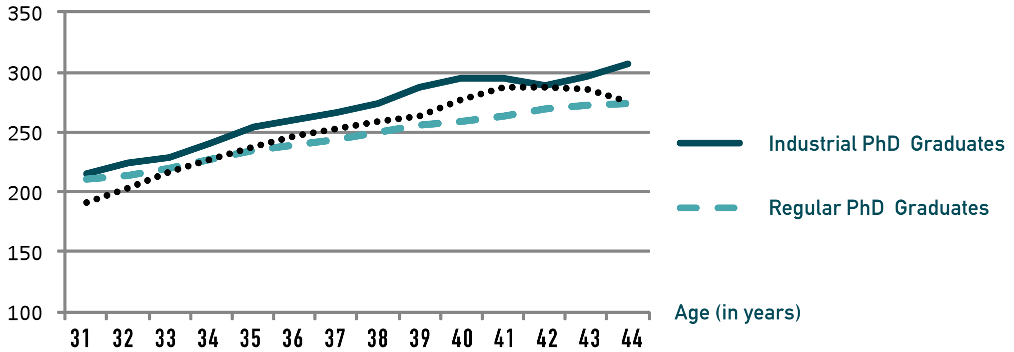
Interestingly, there are a number of Industrial PhDs in medical sciences, yet only relatively few of these individuals had medical science university degrees before taking their PhD.

TABLE 3.2.3: Type of PhD degrees, in percent			
	Industrial PhDs	Regular PhDs	All
Technical sciences	59,5	24,5	25,9
Natural sciences	10,6	22,5	22,0
Other disciplines	3,4	20,8	20,1
Medical sciences	14,3	19,3	19,1
Veterinarian/agricultural	6,3	8,9	8,8
Pharmaceutical sciences	3,2	2,0	2,1
Social sciences	2,7	1,9	2,0

Results

As a first step, we compare average hourly wages of the different groups of individuals under consideration in 2006 and graph the averages as a function of age in FIGURE 3.2.1.

FIGURE 3.2.1: Hourly wage (DKK) in 2006, by Individual age



We find that wages of Industrial PhD graduates are higher than those of both regular PhD graduates and university graduates. Differences are largest in the early and mid-forties. Regular PhDs have lower wages relative to both Industrial PhDs and university graduates.

An obvious explanation of these differences may be sought in different employment patterns of the different groups of employees, with regular PhDs being overrepresented in public sector research institutions, which are generally characterised by lower wages than private sector employers.

Comparisons between Industrial PhDs and regular PhDs

In this section, we compare wages between Industrial and regular PhD graduates by using a linear regression, holding constant a set of (pre-determined) background characteristics (age, gender, etc.).

These comparisons are based on the 430 Industrial and approx. 5,850 regular PhD graduates. We choose a logarithmic specification of the wage variable, implying that regression coefficients are the expected (approximately) percentage-wise changes in the wage when the condition of the associated explanatory variable is fulfilled.

TABLE 3.2.4: Hourly wage of Industrial and regular PhD graduates, linear regression results, dependent variable: log (hourly wage), sample: Industrial and regular PhD graduates (2006)

Variables	Coefficient		Standard error	Coefficient		Standard error
The person is an Industrial PhD graduate	0.090	***	0.014	0.063	***	0.014
The person is female	-0.060	***	0.007	-0.065	***	0.007
The person is an immigrant (or descendant)	0.028		0.027	0.005		0.027
Grade of secondary education diploma (normalised)	0.026	***	0.004	0.019	***	0.003
Age (in years)	0.016	***	0.001	0.021	***	0.001
Additional controls	Secondary education: elective courses (7 categories)			Secondary education: elective courses (7 categories); specific PhD degree (10 categories); age when receiving the PhD degree		
Number of observations	6.283			6.283		

Notes: ***: significant at the 1% level. Heteroscedasticity-consistent standard errors.

The regression results confirm the findings of FIGURE 3.2.1: Industrial PhDs earn approx. (exp(0.090)=1.094) 9 percent higher hourly wages compared to their counterparts who have taken a regular PhD degree. When including additional variables in the regressions (which control for the different compositions of the subjects of the PhD projects and for age differences when obtaining the PhD degree), the difference drops to approx. 6 percent, but remains statistically highly significant.

Here, it might be noted that the result of positive wage income differences is robust when considering gross hourly wages (i.e. total wage income including pensions divided by the number of working hours) or annual income instead of the current wage concept. In the first case, the relevant coefficient dropped to approx. 5 percent (instead of approx. 9 percent). In the second case, when considering annual income without correcting for working hours, the coefficient increased to between 10 (in the specification with additional controls) and 15 percent (in the more simple specification). This indicates that Industrial PhDs register more working hours than regular PhDs.

Comparing the career developments between the two types of PhDs, we first note that 6.3 percent of Industrial PhD graduates are employed in leadership positions, as opposed to 3.9 percent of regular PhD graduates. The formal comparison is by estimating a so-called binary choice model (assuming a logistic distribution). The coefficients of this model are displayed in TABLE 3.2.5.

TABLE 3.2.5: Occupation of Industrial and regular PhD graduates, binary choice (logit) model, sample: Industrial and regular PhD graduates (2006)

	Dependent variable: The person has a leadership position			Dependent variable: The person has a specialist position		
Variables	Coefficient		Standard error	Coefficient		Standard error
The person is an Industrial PhD graduate	1.087	***	0.217	-0.738	***	0.112
The person is female	-0.577	***	0.178	0.268	***	0.064
The person is an immigrant (or descendant)	0.985	**	0.403	-0.344		0.227
Grade of secondary education diploma (normalised)	0.108		0.078	0.077	**	0.033
Age (in years)	0.103	***	0.020	0.010		0.009
Additional controls	Secondary education: elective courses (7 categories)			Secondary education: elective courses (7 categories)		
Number of observations	7,214			7,214		

Notes: ***: significant at the 1% level, ** significant at the 5% level.

The exponents of the model's coefficients equal the increases in the probability that the individual is a leader or a specialist when the logical conditions of the corresponding variables are true. We find that Industrial PhDs are almost three times more likely (exp(1.087)=2.95) to hold a leadership position than regular PhDs when holding constant the set of background characteristics included in the regression.

Industrial PhDs have an approx. (exp(-0.738)=0.48) 50 percent lower probability of being employed in a specialist position than regular PhDs. We conclude that, while regular PhDs are almost entirely employed in specialist positions, Industrial PhDs are more evenly distributed across the different occupational levels.

Comparisons between Industrial PhDs and university graduates

In the following, we compare wages between Industrial PhD graduates and university graduates.
TABLE 3.2.6 summarises the results of the comparison of hourly wages. They suggest that Industrial PhDs earn a wage premium of approx. 7 percent relative to university graduates in similar fields of study while controlling for demographic factors, secondary education grades and course specialisation.

TABLE 3.2.6: Hourly wage of Industrial PhD and university graduates, linear regression results, dependent variable: log (hourly wage), sample: Industrial PhD graduates and university graduates (2006)

Variables	Coefficient		Standard error
The person is an Industrial PhD graduate	0.066	***	0.014
The person is female	-0.129	***	0.009
The person is an immigrant (or descendant)	-0.053		0.034
Grade of secondary education diploma (normalised)	0.043	***	0.005
Age (in years)	0.028	***	0.001
Additional controls	Secondary education: elective courses (7 categories)		
Number of observations	5,246		

Notes: ***: significant at the 1% level. Heteroscedasticity-consistent standard errors.

Again, it may be noted that the result of positive wage income differences is unaffected by considering gross hourly wages (i.e. total wage income including pensions divided by the number of working hours) or annual income instead of the current wage concept. In the first case, the relevant coefficient dropped to 0.056 (instead of 0.066). In the second case, when considering annual income without correcting for working hours, the coefficient increased to 0.11 - again indicating that Industrial PhDs register more working hours than other graduates.

The results of the career development comparisons are found in TABLE 3.2.7. In comparison with university graduates, Industrial PhDs are overrepresented in leadership positions and specialist positions. However, the difference regarding leadership positions is not statistically significant and must be regarded as tentative.

For specialist positions, the coefficient 0.397 corresponds to an approx. 50 percent higher probability that Industrial PhDs are employed in specialist positions than university graduates.

TABLE 3.2.7: Occupation of Industrial PhD graduates and university graduates, binary choice (logit) model, sample: Industrial PhD graduates and university graduates (2006)

	Dependent variable: The person has a leadership position			Dependent variable: The person has a specialist position		
Variables	Coefficient		Standard error	Coefficient		Standard error
The person is an Industrial PhD graduate	0.182		0.217	0.397	***	0.113
The person is female	-0.741	***	0.156	0.105	**	0.051
The person is immigrant (or descendant)	-0.572		0.710	-0.120		0.217
Grade of secondary education diploma (normalised)	0.139	**	0.064	0.151	***	0.025
Age (in years)	0.095	***	0.015	0.072	***	0.006
Additional controls	Secondary education: elective courses (7 categories)			Secondary education: elective courses (7 categories)		
Number of observations	7,465			7,465		

Notes: ***: significant at the 1% level, ** significant at the 5% level.

4 COMPANY LEVEL ANALYSIS >

4.1 Data and methodology of the company level analysis

Data

The data used for the company level analysis is from three sources:

- First, data from DASTI on the participation of companies and individuals in the Industrial PhD Programme.
- Second, information on financial reports that companies above certain size thresholds must file to a public authority.
- Third, information on patenting activity from the European patent office.

The data from DASTI on the participation of companies and individuals in the Industrial PhD Programme contain information on the year an individual was employed as an Industrial PhD student, and in many cases also the employing company's registration number ('cvr-number'), which is filed at the public authorities and which is also available in the other datasets used in this study.

Data on financial reports is from the private information provider company Købmandsstandens Oplysningsbureau, now Experian A/S. This dataset, henceforth denoted as the KOB data, contains information from the financial reports that companies with a certain size and ownership structure must file to the public authorities.

Data on patenting is from the CEBR patent database, which has information on all patent applications at the European Patent Office by at least one applicant residing in Denmark.

The sample

In the original data from DASTI, there are 1,224 Industrial PhD projects in 536 different companies; 47 projects are registered as abandoned. Excluding these projects from the sample (including one project which lacks information on when the project was started) leaves us with 1,177 projects and 514 different companies.

However, it should be noted that in the original sample, the 514 different companies are defined by their names. This number is partly due to registering the same company under slightly different names in the DASTI data.

For the following performance analysis, we have to merge the sample of 1,177 projects in 514 companies with the information from the KOB database.

To accomplish this, we first had to find company registration numbers ('cvr'-numbers) of companies with missing or erroneous registration numbers in the original DASTI data. We managed to find these registration numbers for 509 different companies as defined by their names (hosting 1,161 projects). These 509 different company names in the DASTI data correspond to 445 different companies as defined by their company registration numbers. This is the definition of companies we will use henceforth.

The first Industrial PhD projects were initiated in 1988. Up to 2003, the number of projects initiated each year was relatively stable at approx. 30 to 50. However, in recent years the number of projects initiated per year has increased steadily and is now in the range of 80 to 120 projects. It should be noted that approx. 30 percent of the companies in the sample have hosted more than one project, and that some companies have hosted a considerable number of projects (e.g. more than 20).

1,053 out of the 1,161 projects and 387 out of the 445 companies can be identified in the KOB database. A large share of the attrition is related to companies which have either been established too recently to be covered by the KOB database or closed down before the KOB database assumed full coverage.

For 383 companies, there is financial report information in the KOB data. These companies are observed on average for 15.6 years, which implies that there are a total of 5,018 annual financial reports for companies that have hosted at least one Industrial PhD project. However, it should be noted that any potential bottom-line effects of Industrial PhD projects may take a couple of years to materialise, and that a considerable share of Industrial PhD projects was initiated at the end of the observation period.

Of the 383 companies in the KOB data, 72 companies are not observed after first initiating an Industrial PhD project. These obviously cannot be used for the following analysis, leaving us with 311 different companies have hosted a total of 851 Industrial PhD projects. Out of these, 195 companies have hosted only one Industrial PhD project, 48 percent have hosted two projects, 27 companies three projects, 9 companies four projects, and 32 (approx. 10 percent) companies more than five projects. There are also a few companies which have hosted more than 20 projects.

The companies with many projects are typically large companies for which it is difficult, if not impossible, to find similar companies for the comparisons in the statistical analyses to follow. Also, the statistical model which is preferred by the precision of its estimates requires fixing a year before a company first participates in the Industrial PhD Programme. For companies with many projects, this year is not well-defined, and the year before hosting the first Industrial PhD is often before the KOB database assumes full coverage.

Accordingly, we will only consider companies that have hosted a maximum of three projects for the following analysis. These represent approx. 85 percent of all companies participating in the Industrial PhD Programme, which leaves us with 270 companies for the company level analysis.

Of the 270 companies, approx. 120 are observed five years before first initiating an Industrial PhD project, approx. 160 are observed five years after, and 86 are observed ten years after first initiating a project.[9]

The characteristics of the companies in the sample used for analysis are

[9] However, it should be noted that missing information for a number of observations means that the number of records which can be used for the analysis is reduced. For example, total factor productivity figures are available for 91 companies five years after first initiating an Industrial PhD project, and for 46 companies ten years after first initiating a project.

TABLE 4.2.1: Descriptive statistics of the matched treatment-control samples

		All companies with a maximum of three projects			High-quality matches		
		Companies that have hosted at least one Industrial PhD project	Control companies	All companies	Companies that have hosted at least one Industrial PhD project	Control companies	All companies
Number of companies		270	539	809	129	283	412
	Total factor productivity	-0.056	-0.006	-0.023	0.090	0.016	0.039
	Gross profit per employee (DKK1,000)	1529.4	689.5	971.9	445.5	466.6	460.0
	Patent applications	2.4	0.7	1.3	0.8	0.3	0.5
	Number of employees	520.0	212.0	314.8	28.8	31.8	30.9
	Gross profit (DKK1,000)	651460.8	154181.9	320482.4	14828.6	15841.2	15522.5
	Total assets (DKK1,000)	18800000	951008	6894683	22515.98	22661.74	22616.1
	Establishment year	1978.8	1977.8	1978.2	1988.6	1988.2	1988.4
Industries							
	Business services	18.52	18.55	18.54	26.87	24.09	24.94
	Research and development	9.26	9.09	9.15	14.93	12.54	13.27
	IT	8.89	8.91	8.90	11.94	11.22	11.44
	Medical equipment, instruments manufacturing	8.52	8.53	8.53	7.46	7.92	7.78
	Finance	8.52	8.53	8.53	5.97	6.60	6.41
	Wholesale trade	7.78	7.79	7.79	5.97	4.95	5.26
	Chemicals, pharmaceuticals	4.81	4.82	4.82	3.73	3.96	3.89
	Food production	4.44	4.45	4.45	0.00	2.97	2.06
	Manufacturing	3.33	3.34	3.34	2.99	2.31	2.52
	Other	25.93	25.97	25.96	0.00	0.00	0.00
Zip-codes							
	1000-1999	11.85	11.13	11.37	12.69	11.55	11.90
	2000-2999	41.85	39.89	40.54	44.03	38.61	40.27
	3000-3999	10.00	9.83	9.89	11.19	10.89	10.98
	4000-4999	4.07	4.27	4.20	3.73	4.62	4.35
	5000-5999	7.41	7.98	7.79	8.21	6.93	7.32
	6000-6999	3.33	5.38	4.70	1.49	4.29	3.43
	7000-7999	4.81	3.53	3.96	0.75	1.98	1.60
	8000-8999	9.63	11.13	10.63	8.96	10.23	9.84
	9000-9999	7.04	6.86	6.92	5.22	4.29	4.58

Unsurprisingly, we find that Industrial PhDs are typically hosted by companies in knowledge-intensive industries. Also, hosting companies are geographically concentrated in the Copenhagen area (zip-codes below 3000).

Companies hosting Industrial PhD projects are, on average, relatively large companies with sometimes very high capital intensities (which is mostly due to the presence of large financial sector companies).

Methodology of the company level analysis

Our statistical model compares two groups of companies:
(a) companies that have hosted at least one Industrial PhD project, and
(b) companies that have not hosted any Industrial PhD projects.

In accordance with the academic project evaluation literature, the group of companies which have hosted Industrial PhD projects will henceforth be called the ‘treatment group’, while the comparison group of companies which have not hosted any Industrial PhD projects will be denoted as the ‘control group’.

When interpreting the results of the statistical comparisons, one must take into account the fact that it is not possible to include all relevant factors in the models because they are unobservable in the data. Examples include different kinds of company competences and other immeasurable company characteristics.

This implies that interpreting any systematic treatment-control differences in company performance developments as genuine causal effects of hosting an Industrial PhD project will have to rest on an ‘all-else-equal’ assumption, i.e. the assumption that factors omitted from the model are either irrelevant or, on average, equal for treatments and controls.

To maximise the validity of this ‘all-else-equal’ assumption, we identify the control group using a matching procedure which ensures that we compare the treatment group companies with a control group of highly similar companies.

The identification procedure is described in greater detail in Appendix 1. Here, it may be sufficient to note that in the analysis to follow, we will compare developments in the success parameters over time of two groups of companies highly similar in a number of observable characteristics.

Of interest in the following analysis is whether treatment group companies experience more positive developments in the success parameters in association with hosting Industrial PhD projects compared to control group companies.

The modelling setup was chosen to generate the most precise estimates possible. However, it should be noted that the associated before/after comparisons imply that this procedure is only applicable to analysing companies that have hosted one or very few projects, as otherwise the timing issue cannot be resolved.

As a compromise between the precision of the before/after time period definition and having a sufficient number of observations for the analysis, we consider companies with a maximum of three Industrial PhD projects. As noted earlier, these companies represent approx. 85 percent of the participating companies.

The year that separates a company's pre-participation period from its post-participation period will be denoted as "year 0" or the "base year". For companies hosting an Industrial PhD project, year 0 is defined as the year before initiating the first Industrial PhD project. For a company in the control group, year 0 is the year in which it most resembled one of the project hosting companies in its base year.

Using this method, we can measure participating companies' developments in the success parameters before and after their base year - the year before initiating the first Industrial PhD project - and compare these developments to the developments of the control group companies.

4.2 Results of the company level analysis

In the following sections, the results of the company level analysis will be described. Here, two introductory remarks should be made:

Firstly, it must be assumed that it is practically impossible to isolate any performance effects of hosting Industrial PhD projects on large companies, as any contribution of an Industrial PhD project on aggregate company performance would be small relative to the companies' considerable heterogeneity in the success measures. For this reason, we will also present results for an alternative sample where companies with more than 300 employees or total assets of at least DKK 100 million in year 0 are not considered. This sample will be denoted the 'sample of small companies'.

Secondly, it proved to be difficult to find highly similar control companies for a number of treatment companies. For this reason, we also consider a separate sample of companies with less than 300 employees and total assets of less than DKK 100 million where these low-quality matches are excluded. This results in a sample of highly similar treatment and control group companies, denoted as the sample of 'high-quality matches'.

Before turning to the comparisons of the company performance parameters, we will address the question of how successful the matching procedure is in finding highly similar groups of treatment and control companies. Turning back to TABLE 4.2.1, which is a snapshot of the companies in year 0, we can compare the observable characteristics of the treatment and control group companies – both for the sample of all companies with a maximum of three Industrial PhD projects and their corresponding control companies, and for the sample of high quality matches.[10]

[10] Note that the sampling procedure implies that the base years of the two groups of companies is distributed highly similarly over time.

While industry and geographical distributions are almost identical for treatment and control group companies in the two samples (implied by the matching procedure), some of the very large Industrial PhD companies in the sample of all companies lack counterparts in the control group. In this sample of all companies, treatment group companies have a lower total factor productivity and a higher gross profit (which is consistent with a higher capital intensity) than the control group companies. However, the large heterogeneity in these variables implies that these differences are not statistically significant.

Companies in the high quality match sample are on average considerably smaller, younger and, of course, generally more similar in their observable characteristics.

We conclude that it was possible to find highly similar matches in terms of geographic location and company age. For the sample of high-quality matches, controls are also highly similar in company size.

Patenting activity

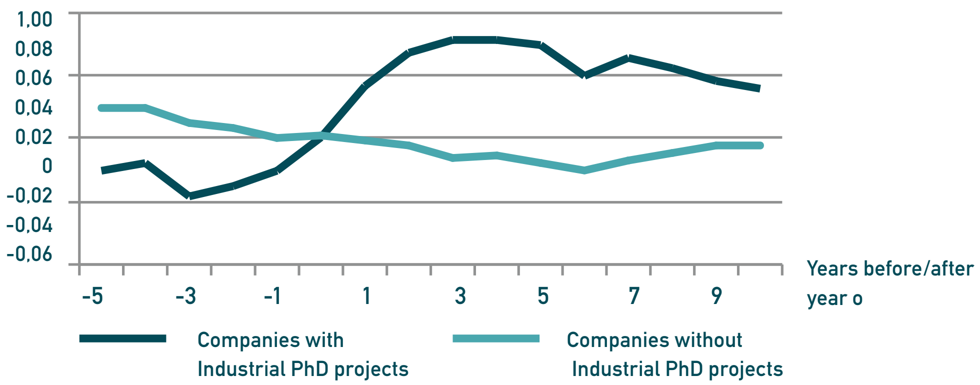
Patenting activity is measured by the company's number of patent applications per year.[11]

To isolate any Industrial PhD programme participation effects, we calculate for every company and year the difference between the number of patent applications in the given year and the number of patent applications filed in year 0.

FIGURES 4.2.1-3 display developments of these differences, i.e. current patenting activity relative to activity in year 0 for treatment and control group companies, respectively.

We find large movements over time for companies that host Industrial PhD projects relative to companies in the control group. This is likely to be a result of generally higher absolute patenting activity in treatment companies.

FIGURE 4.2.1: Number of patent applications, all companies



[11] An alternative measure would have been to consider granted patents. However, the long patent approval process renders it difficult to associate this variable to current innovation output.

FIGURE 4.2.2: Number of patent applications, small companies

Average number of patent applications per company relative to year 0

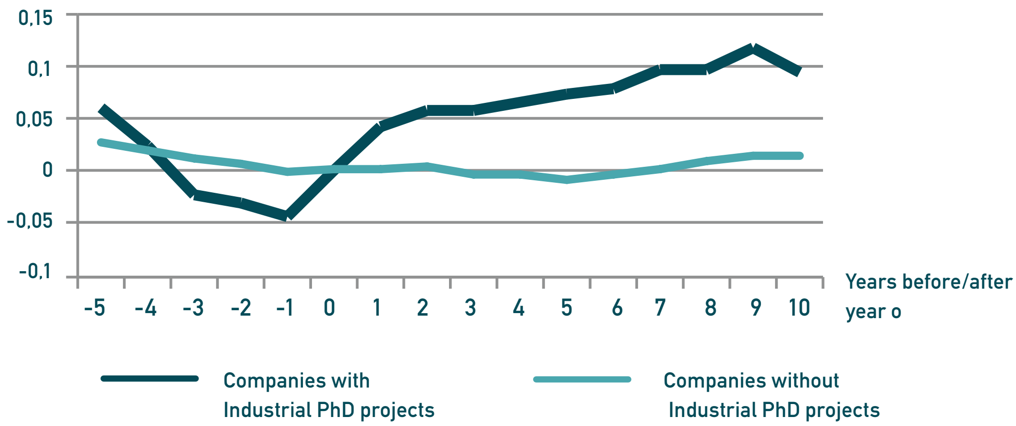
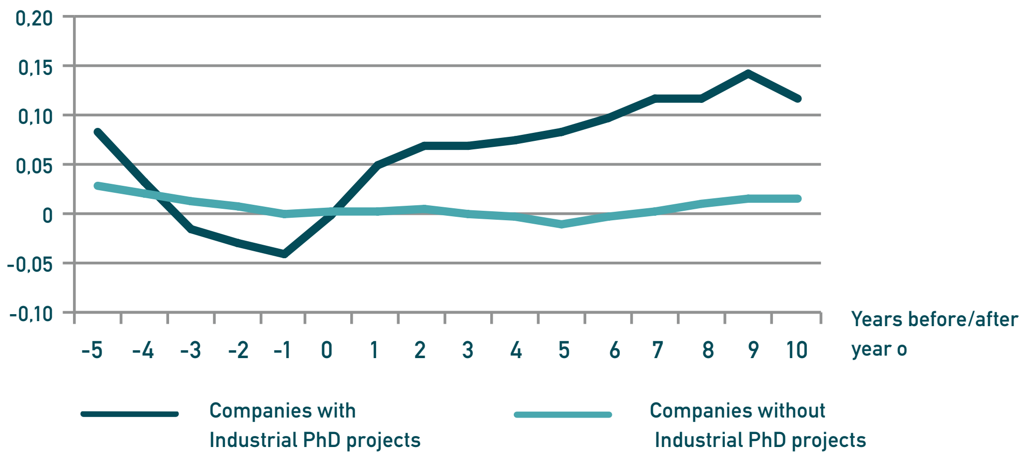


FIGURE 2: Number of patent applications, high-quality matches

Average number of patent applications per company, change relative to year before first initiating an Industrial PhD project



All graphs indicate that after year 0, the developments over time for treatment companies are equal to or larger than developments for control companies, indicating greater increases in patenting activity for the group of treatment companies compared to the group of control companies.

One could note that there are also differences between pre-base year trends in patenting activity depending on the sample under consideration, indicating the difficulties of finding control companies with patenting activities similar to the companies hiring Industrial PhDs.

>

Whether or not one is willing to interpret the graphs as evidence of positive effects of hosting Industrial PhD projects depends on one's underlying assumptions. E.g. in FIGURE 4.2.1, there is a positive trend of increasing patenting activity before hosting the first Industrial PhD project, but not in the years after. But over a longer time horizon, activity is higher after year 0 than before. So the interpretation of the results depends on whether one assumes that:
(a) trends would continue in the absence of programme participation, or,
(b) activity would stay at the same level in the long run in the absence of the programme.

The estimates of the statistical model presented below will be based on a pre-participation period specified as the five years up to year zero, and the post-participation period as the ten years after year zero. Obviously, the lengths of these time periods are computed are arbitrary, and the robustness of later results when choosing different before/after time intervals needs to be checked in the numerical analysis.

A look at the raw data reveals that participant firms in the sample of high-quality matches apply for on average 0.07 patents per year before year zero, and 0.18 after year zero (i.e., an increase of 0.11). Control firms have almost the same patenting activity both before and after year zero. Under the assumption that both groups of firms would have experienced the same developments in their patenting activity in the absence of the programme, the programme increases patenting activity with 0.11 patent applications per year.

To address the robustness of the graphs' suggestions and to quantify the strength of these associations in the data, we apply a model that estimates the expected percentage-point changes in the number of patent applications in a given year depending on whether the company is a treatment or a control company, and on whether the year under consideration is before or after the base year.

The results of this model are presented in TABLE 4.2.2. Of particular interest are the coefficients for the variable "The observation is after the base year and belongs to an Industrial PhD company". Under the assumption that patenting of treatments and controls would develop in similar ways in the absence of the programme, this variable identifies the genuine causal effect of hosting an Industrial PhD project on patenting activity.

TABLE 4.2.2: Count data regression results, dependent variable: number of patent applications in a given year. The table presents exponentiated coefficients, i.e. multiples of the number of patent applications when the logical conditions of the associated variable are fulfilled.

		Sample: all compa-nies with a maximum of three Industrial PhD projects		Sample: small companies		Sample: high-quality matches	
Variable							
	The observation is after year 0	0,88		0,76		0,84	
	The observation belongs to an Industrial PhD company	4,13	***	4,17	***	4,36	***
The observation is after year 0 and belongs to an Industrial PhD company		1,70	**	2,19	**	1,94	*
	Constant term	0,06	***	0,04	***	0,04	***

Notes: ***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level. All regressions based on STATA Corp.'s 'xtpoisson' routine.

Findings of the statistical analysis of patenting activity can be summarised as follows: We find positive potential effects of hosting an Industrial PhD project for the sample of all companies with a maximum of three Industrial PhD projects. According to the estimates, hosting an Industrial PhD almost doubles (1.70) the number of patents per year in the years after year 0.

For the other samples, associations between hosting Industrial PhD projects and changes in patenting activity are also positive, and stay significant the ten-percent significance level also for the considerable reduced sample of high-quality matches.

In sum, one can conclude that there is evidence of positive associations between hosting Industrial PhD projects and changes in patenting activity in the data.[12]

[12] These relationships were robust when changing the lengths of the before- and after-base year periods considered for the estimations. Also, computing average numbers of patents of both participants and controls both before and after year zero, and estimating a linear model of the pre-post base-year differences revealed very similar (and also statistically significant) results.

Gross profit growth

The analyses of gross profit and TFP in the next subsection follow the same blueprint as the previous look at patenting activity. Recall that gross profit is the surplus of annual revenues over costs (excluding wages), and accordingly measures the value creation of a company in a given year.

First, for every year we calculate the difference between the year's gross profit and the gross profit in the base year (the year before the company first initiated an Industrial PhD project). Next, we calculate the average of these differences for both the group of treatment companies (which have hosted Industrial PhD projects) and the group of control companies (which have not hosted any Industrial PhD project).

FIGURES 4.2.4-6 show these averages for the treatment and control companies for the three different samples. They suggest that companies which host Industrial PhD projects are characterised by high growth in gross profit. While FIGURE 4.2.4, which compares all sampled companies both with and without Industrial PhD projects, show a decrease in the growth trend in association with hosting the first Industrial PhD project, FIGURE 4.2.5 and FIGURE 4.2.6, respectively comparing small companies and high quality matches, show a consistent gross profit growth which has no equivalent in the corresponding control group's gross profit growth pattern.

FIGURE 4.2.4: Gross profit developments (in DKK1,000), all companies

Average values relative to year 0

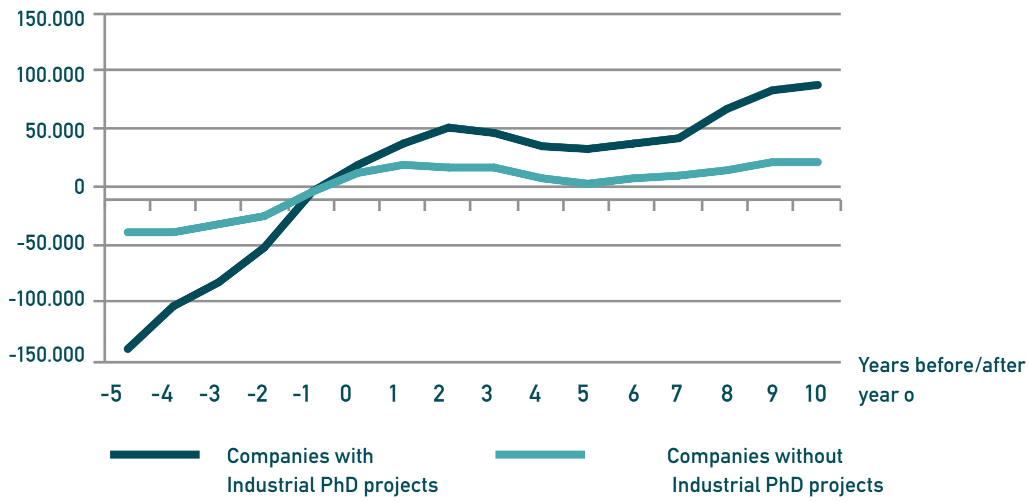


FIGURE 4.2.5: Gross profit developments (in DKK1,000), small companies

Average values relative to year 0

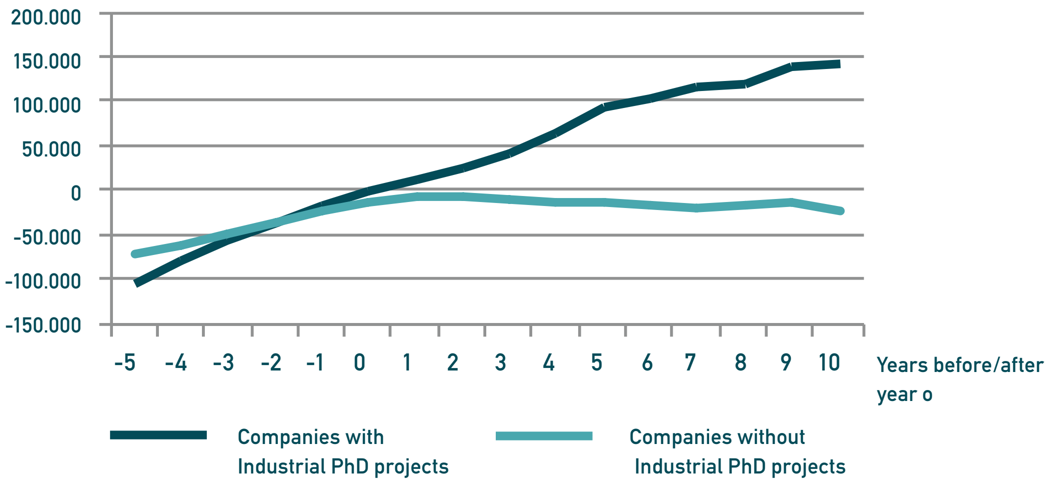
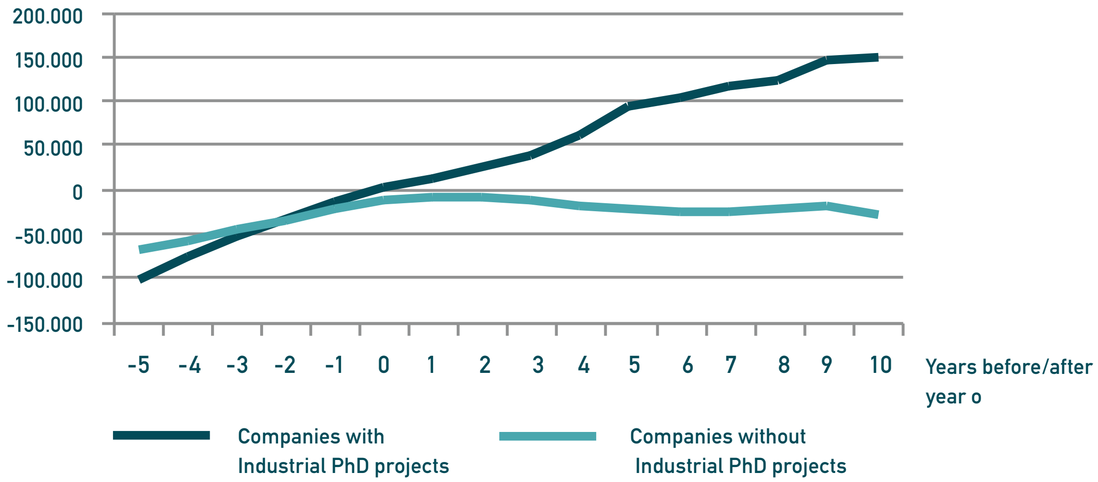


FIGURE 4.2.6: Gross profit developments (in DKK1,000), high-quality matches

Average values relative to year 0



Thus, assuming that treatment group companies in the absence of the programme would experience similar gross profit growth (or in this case: a similar decline in growth) as control group companies, the vertical distance between the graphs shows a considerable genuine (causal) effect on the gross profit growth of companies hosting Industrial PhD projects.

We turn now to formally estimating before/after year 0 differences in gross profit growth for treatment and control groups respectively. [13] Accordingly, we divide

[13] We consider before/after differences in growth rather than levels, since gross profit levels show clear time trends which need to be taken into consideration in the estimations to avoid generating biased estimates.

each company's observation period into two periods: one period before year 0, and one period after year 0. For every company and for both two periods, gross profit growth is measured by the average of the annual (absolute) increases in gross profit.

We can now compare these averages both over time and between treatment and control group companies. In the statistical model, we use the same variables as in the count data regression used of the patenting analysis as right-hand-side variables.

Hence, the difference between the developments in gross profit growth before/after year 0 for treatment group companies and the gross profit growth developments before/after year 0 for control group companies is measured by the coefficient associated with the variable: *"The observation is after the base year and belongs to an Industrial PhD company"* [14].

The results of this comparison, which is again carried out by using a simple linear regression model, are summarised in TABLE 4.2.3. The table shows the results for high-quality matches, i.e. the treatment and control group companies most similar to each other with regard to their observable characteristics, and for which the comparison accordingly has the highest validity.

TABLE 4.2.3: Linear regression results, dependent variable: annual increase in gross profit (in DKK1,000, in prices of 2007), sample: high-quality matches.

		Observation period: three years before to five years after year 0			Observation period: three years before to ten years after year 0		
Variable		Coefficient		Standard error	Coefficient		Standard error
	The observation is after year 0	-1792,35	**	764,36	-1422,68	*	797,83
	The observation belongs to an Industrial PhD company	-458,33		905,19	-458,33		906,09
The observation is after year 0 and belongs to an Industrial PhD company		2267,23	*	1259,42	1458,82		1457,88
	Constant term	1488,27	**	607,25	1488,27	**	607,86
	Number of observations	381			321		

Notes: ***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level. Estimated with heteroscedasticity-consistent standard errors.

[14] E.g., if the increase in annual gross profit of treatment companies is on average DKK 5m before the base year and DKK 7m after year 0, and if gross profit for control firms increases on average DKK 3m before and DKK 4m after year 0, the coefficient associated with "The observation is after the base year and belongs to an Industrial PhD company", measured in DKK, is equal to (7m-5m)-(4m-3m)= 1m.

In the first model, which compares growth trends in the three-year period before and the five-year period after year 0, the coefficient of *"The observation is after year 0"* (-1,792.35) suggests that gross profit growth has slowed down by almost DKK 2m per year.

But the estimate of the coefficient for the variable "*The observation is after the base year and belongs to an Industrial PhD company"* of 2,267.23 implies that gross profit growth of Industrial PhD companies maintains its positive trend. So, for Industrial PhD companies, growth after first initiating an Industrial PhD project is approx. DKK 2m higher per year than would otherwise be expected if they had experienced a similar decline in gross profit growth as the control group companies.

So the approx. DKK 2m growth difference per year, implying an additional gross profit of (2+4+6+8+10) DKK 30m in the first five years of programme participation, is the genuine causal effect of programme participation, assuming that Industrial PhD companies' growth in gross profit would otherwise have followed the exact same pattern of the control companies if they had not participated.

It becomes clear that the programme might be considered successful even if only a part of this difference is because of a genuine causal effect of the Industrial PhD Programme.

When we compare the growth patterns of the two groups of companies between both the three-year time period before and the ten-year time period after year 0, the difference still suggests higher growth for participating companies, but becomes statistically insignificant (i.e. it becomes more likely that the finding is coincidental).

Total factor productivity
For this analysis, total factor productivity (TFP) was calculated on an annual basis for all companies in the entire KOB database in the given year.

Total factor productivity is gross profit 'corrected for' the number of employees and total assets. It is calculated as the residuals of a Cobb-Douglas-production function regression. In other words, TFP is the share of the company's value creation which cannot be explained by its number of employees or its capital stock.

Thus defined, TFP approximates the percentage-wise deviation in gross profit from the gross profit that we would have expected to observe, given the company's number of employees and its stock of assets.

For the analysis, we first take a look at the developments using a graphical depiction of the data. FIGURES 4.2.7-9 summarise.

FIGURE 4.2.7: Total factor productivity developments, all companies.

Average values relative to year 0

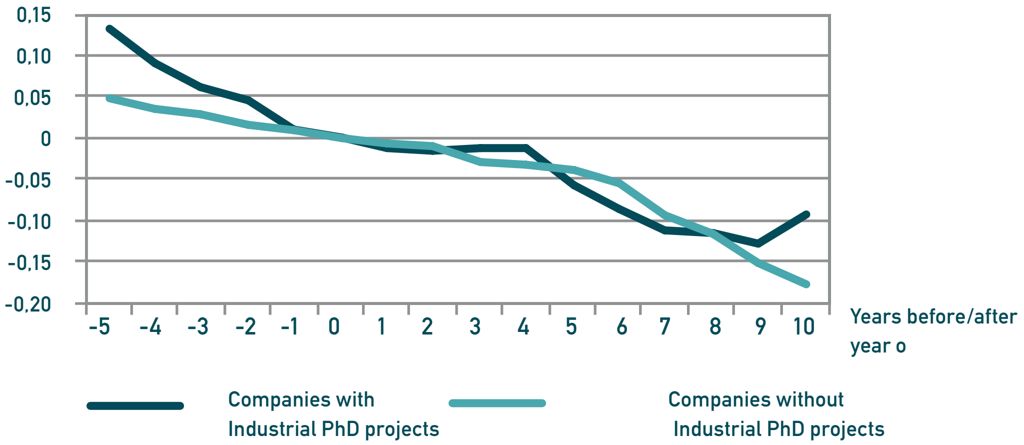


FIGURE 4.2.8: Total factor productivity developments, small companies

Average values relative to year 0

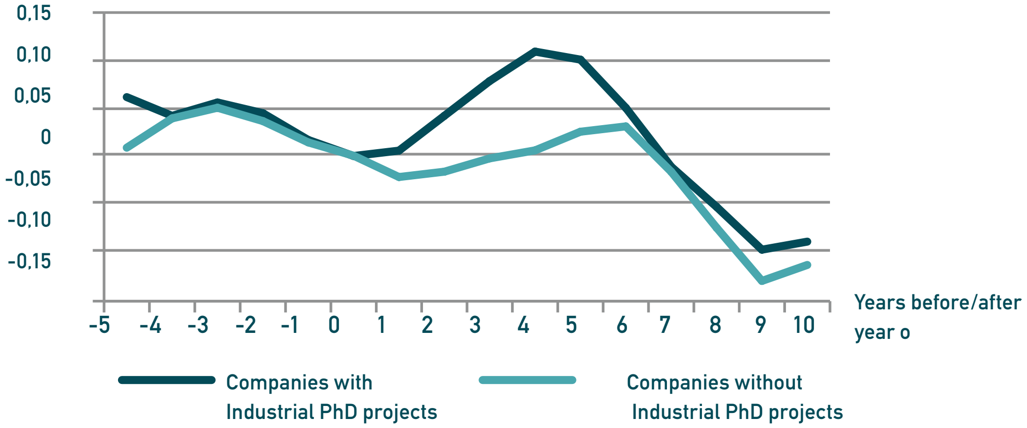
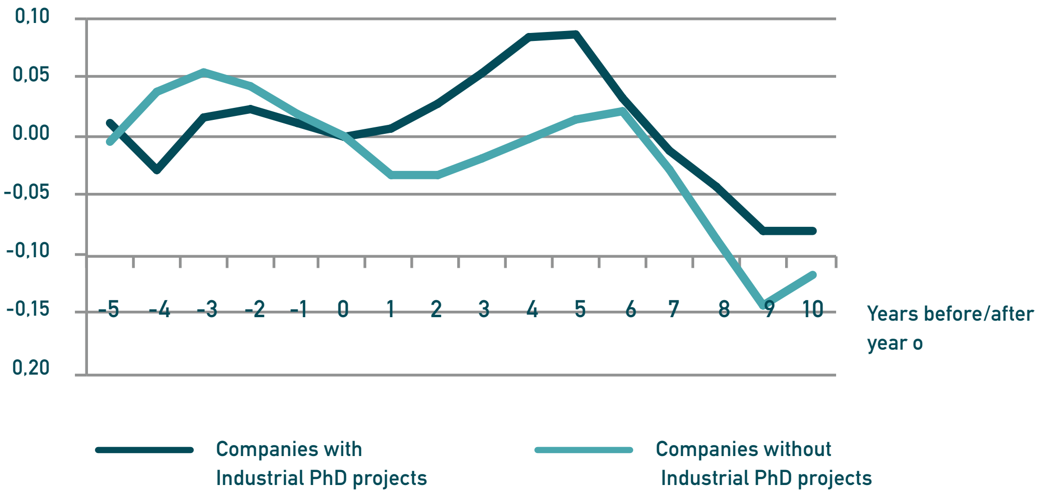


FIGURE 4.2.9: Total factor productivity developments, high-quality matches

Average values relative to year 0



The figures illustrate that developments are very different depending on whether or not large companies are excluded from the sample observed.

While there is a negative trend in TFP for the sample of all companies, there are no such trends for the subsamples. The erratic movements in the graphs (in spite of smoothing) suggest large heterogeneity in TFP over time and between companies.

For the subsamples, which are unaffected by the presence of very large companies, TFP is between 5 to 10 percentage points higher approx. two to six years after year 0 in the subsamples.
Again, we qualify the suggestions of the graphs by use of linear regression, the results of which are depicted in TABLE 4.2.4.

TABLE 4.2.4: Linear regression results, dependent variable: (TFP in a given year) - (TFP in year 0)

	Sample: all companies with a maximum of three Industrial PhD projects			Sample: small companies		Sample: high-quality matches		
Variable	Coefficient		Standard error	Coefficient	Standard error	Coefficient		Standard error
The observation is after year 0	-0,084	***	0,023	-0,050	0,036	-0,066	*	0,038
The observation belongs to an Industrial PhD company	0,057		0,028	0,008	0,043	-0,027		0,043
The observation is after year 0 and belongs to an Industrial PhD company	-0,002		0,040	0,042	0,064	0,068		0,067
Constant term	0,003		0,014	0,013	0,019	0,019		0,020

Notes: ***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.
Estimated with heteroscedasticity-consistent standard errors

We see the negative TFP trends of companies hosting Industrial PhD projects and their counterparts in the high-quality match control group corroborated by the negative coefficients associated with the variable *"the observation is after year 0"*.

Also, TFP has increased more (or decreased less) in the treatment group companies compared to the control group the samples of small companies and that of the high-quality matches.

This is indicated by the positive coefficients of the variable *"the observation is after year 0 and belongs to an Industrial PhD company"*, which, for high-quality matches, show that companies which have hosted Industrial PhD projects have on average approx. 7 percentage points higher TFP than would otherwise be expected if they had experienced a TFP development similar to the control group companies.

Under the assumption that treatment group companies would experience TFP developments similar to those for control group companies in the absence of initiating Industrial PhD projects, this 7 percentage point difference is the most qualified assumption of the Industrial PhD Programme's causal total factor productivity effect. However, although positive, the TFP differences between treatment and control groups are too small compared to the large variations in TFP to interpret them as statistically significant, and must accordingly be interpreted tentatively. In conclusion, one cannot claim any strong association between hosting Industrial PhD projects and TFP development.[15]

Employment growth

We conclude the company level analysis by taking a look at employment growth. The finding of high growth in gross profit but not in total factor productivity might be an indication that companies hosting Industrial Phd projects are high-growth companies. This is strongly supported by a closer look at the data, illustrated by FIGURE 4.2.10, with companies hosting Industrial PhD projects being characterised by high growth in their number of employees both before and after first initiating a project.

To establish the statistical significance of this result, we formally test the growth difference by means of linear regression, the results of which (for high-quality matches) are presented in TABLE 4.2.5. The results of these regressions suggest that companies participating in the programme sustain an annual employment growth of approximately (-3.48-1.33+3.44+2.95=) 1.58 employees per year in the first five years after first initiating an Industrial PhD project, while companies in the control group decrease their number of employees by approximately -(2.95-3.48=) 0.5 employees per year. Qualitatively, this finding is independent of whether one follows the firms for five or ten years after the base year, and is statistically highly significant.

[15] This finding was robust to changes of the lengths of the time periods before and after the base year which were considered in the regressions. The findings was also robust to changing the regression model, e.g. using each firm's average total factor productivity in the time periods before and after the base year as the dependent variable, or using different specifications of the production function which was employed for the calculation of TFP.

FIGURE 4.2.10: Number of employees developments, high-quality matches

Average values relative to year 0

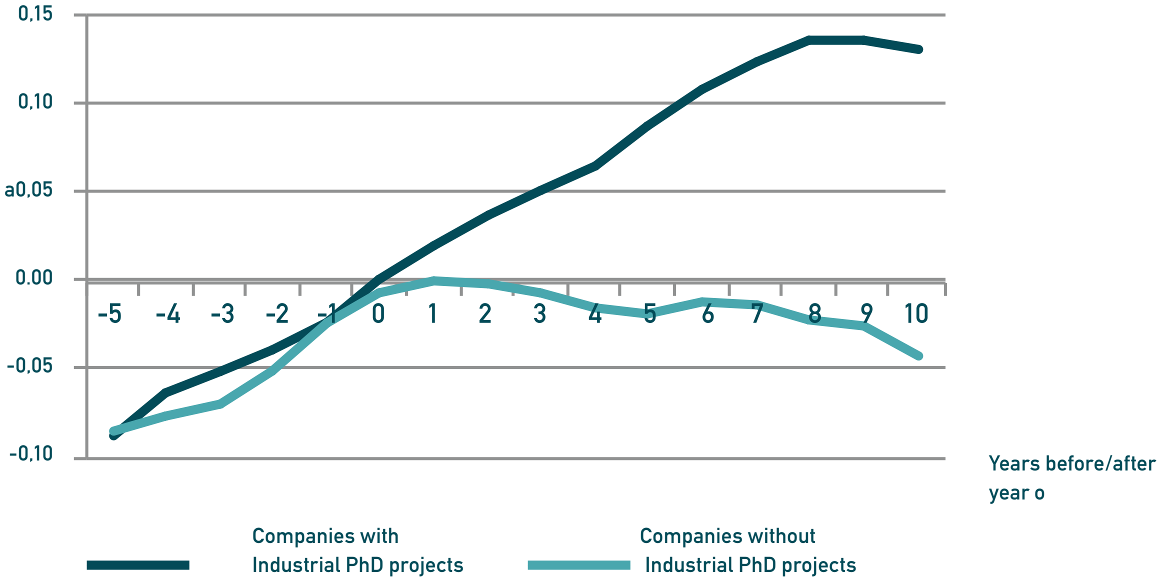


TABLE 4.2.5: Linear regression results, dependent variable: annual increase in number of employees. Sample: high-quality matches

	Observation period: three years before to five years after year 0			Observation period: three years before to ten years after year 0		
Variable	Coefficient		Standard error	Coefficient		Standard error
The observation is after year 0	-3.48	***	0.77	-3.13	***	0.73
The observation belongs to an Industrial PhD company	-1.33		1.00	-1.33		1.00
The observation is after year 0 and belongs to an Industrial PhD company	3.44	***	1.18	2.73	**	1.19
Constant term	2.95	***	0.65	2.95	***	0.65
Number of observations	349			267		
Notes: ***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level. Estimated with heteroscedasticity-consistent standard errors.						

Notes: ***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.
Estimated with heteroscedasticity-consistent standard errors

5 SUMMARY AND CONCLUSIONS >

This analysis considers approx. 430 individuals and 270 companies which have participated in the Industrial PhD Programme and can be found in register data. On the individual level, we compare wage income and the occupations of Industrial PhD graduates with regular PhDs and individuals who have a university degree (and who are similar in terms of their fields of study, gender, etc.).

In the analysis, we take into account a set of demographic background characteristics, like age and gender, but also the average grade of the school-leaving examination, which to some extent controls for individual abilities.

On the company level, we analyse developments across four success parameters: the number of patents, gross profit and employment growth and total factor productivity. For a sample of companies which have hosted a maximum of three Industrial PhD projects before 2009, we identify a control group of highly similar companies which have not hosted any Industrial PhD projects, and compare developments in the success parameters between these two groups. Under identifying assumptions, these models isolate the causal impact of the programme on companies hosting Industrial PhD projects.

The results of the analysis can be summarised as follows: Industrial PhD earn approx. 7-10 percent higher wages than both regular PhDs and university graduates. They are more likely to be found at the top levels of their organisations' hierarchies compared to regular PhDs and more likely to be found in positions requiring high-level specialist knowledge than regular university graduates. Companies which host Industrial PhD projects see on average increasing patenting activity in association with hosting the projects. They are characterised by high growth in gross profit (value creation) and employment.

The comparison with a control group of highly similar control companies suggests that companies hosting Industrial PhD projects would have considerably less positive gross profit and employment developments if they did not participate in the programme.

We cannot find robust differences in total factor productivity developments between companies which have hosted Industrial PhD projects and companies which have not. This finding might be due to firm growth being negatively associated with productivity developments. [16] The relative high wages of Industrial PhD graduates, on the other hand, indicate that they have high individual productivity.

Summing up, earlier studies which found that Industrial PhDs are characterised by positive labour market outcomes have been corroborated. Findings on the company level indicate that the Danish Industrial PhD Programme also has positive effects for participating companies in terms of firm growth and patenting activity.

---

[16] This would be the case if there are decreasing returns to labour, which is one of economic theory's most standard arguments. Empirical support for this argument can be found in: Bingley, P., Westergaard-Nielsen N., 2004, "Personnel policy and profit." Journal of Business Research 2004; 57: 557-563.

6 APPENDIX 1: SELECTION OF CONTROLS

The KOB dataset is a panel dataset which has repeated observations for most of the companies - one for each annual account filed to the authorities. So for every company, there are typically multiple company-year observations (where a company-year observation refers to a record, i.e. a data-point of a given company in a given year). In the following, we will use the expression 'control observation' to describe a single company-year observation (record) of a control company.

Control companies are selected in the year in which they most closely resemble one of the companies participating in the programme, based on the participating company's characteristics in the year before hosting its first Industrial PhD project.

Note that the similarity between participating companies and potential control companies is determined by (a) the companies' region, size, age and industry, and (b) the expected probability of participation, which is derived as follows:

We run an auxiliary regression on the universe of approx. 370,000 company-year observations in KOB in the period from 1994 to 2008 which roughly resemble the group of participants (for example, we do not consider industries in which there is no single participating company).

The auxiliary regression is formulated as a simple probit model where the dependent variable is initiating an Industrial PhD project the following year, and company size, industry, region, productivity, total assets and time period as the model's right-hand-side variables. The regression's pseudo R squared, which is a measure of the model's goodness-of-fit, is 0.29, which we consider to be high.

The probit regression predicts how likely programme participation is for a given company. This allows us to find pairs or groups of companies for which this predicted probability is very similar. For two companies, A and B, with similar participation probability, the fact of company A participating and company B not participating can accordingly be interpreted as coincidental.

Under this interpretation, the identification setup resembles an experiment where programme participation is random, which would allow systematic differences in outcome variables between participants and controls to be interpreted as the programme's causal effect on participating companies.

Yet, even companies with similar predicted participation probabilities can be quite different, and to avoid systematic differences in industry affiliation, size, etc. between participants and controls, we also require that a number of observable characteristics are equal for a given participant and its matched control company(s).

To do this, we divide the total number of company-year observations into groups with the same industry affiliation, same geographic location, of similar size and observed in the same year.

>

For each participating company, we select the company-year observation of a non-participating company within the same group and with a participation probability closest to the participating company's participation probability. This selected company-year observation defines the participating company's control company, and the control company's 'base year' (or 'year 0') – which is the year in which it is most similar to one of the participating companies in its base year, and in which it is selected as a control company. For each of the control companies found by this procedure, the base year forms the basis for comparisons of given success parameters over time.

By repeating the matching procedure, we can find an arbitrary number of control observations for each participant. Here, a greater number of control observations increases the robustness of later results. However, increasing this number also makes it increasingly difficult to find highly similar control observations for some of the participants.

As a compromise between these two considerations, we choose to find two control observations (company-year observations of non-participants) for each participating company. The selection of the two control observations per participating company is made in two rounds. In each of the rounds we select one control observation for each participating company.

In the first round, we find 270 control observations of non-participating companies. In the second, we find another 269 control observations of non-participating companies (the reason for only 269 instead of 270 is that in a single case, one company-year observation is chosen as a control observation for two participants).

In each of the two rounds, we first require that many factors are highly similar when selecting control observations. This leaves a number of participating companies for which no control observations could be found. In subsequent steps, we reduce the number of factors and start choosing control observations which are increasingly less similar, until each round has identified one control observation for every participating company.

When control observations are equal in terms of industry (when distinguishing between at least 36 different categories), number of employees (at least 11 different categories), gross profit (at least 7 categories), time period (at least 7 different categories) and company age (at least 3 different categories), they are regarded as 'high-quality matches' in the analysis.

Note that in each of the rounds, we select only one control observation per participating company. This does not rule out selecting different control observations (belonging to different years) of the same control company. This implies that there are a number of control observations that occur more than once in the data forming the basis