Supplementary note on company participation analysis

This supplementary note is made to clarify some of the more technical issues when performing the company participation analysis. The note will include some of the more technical details regarding the sample, the control group, the selected nace categories and the method used: propensity score matching.

The full sample vs. the final matching sample

Before we can begin the actual matching analysis, some assumptions are needed. We start by illustrating how we get to the 110 companies in the final propensity score matching. From the innovation Denmark database we have a total of 792 unique participating companies. They participate between the years 2003-2014. The information regarding the FP6/FP7 companies are retrieved from eCORDA, the official database from the European Commission. We lose the greatest number of observations by merging with registry data. The registry data are only updated till 2012, meaning that two years of the FP6/FP7 data are disregarded due to this. Also especially data on the companies' educational composition result in fewer observations. However it is one of the criteria, that the companies in the control group have to employ at least one Master or PhD educated person. This is one indicator, describing companies that are likely to participate in either FP6/FP7 but for some reason did not. All the steps are illustrated below in the table.

Table 0.1

The effective number of participation observations in the analysis

	# participating companies	Time	Revenue	FTE	PhD share	Export intensity
Innovation Denmark database	792	2003-2014	-	-	-	-
Matched with registerdata	523	2003-2012	1,070,320,832	1,586	6%	39%
Below 1.000 employees	469	2003-2012	249,312	148	6%	36%
Minimum 1 Master/PhD is employed	400	2003-2012	282,124	164	7%	38%
Removing outliers	314	2003-2012	228,489	137	8%	39%
# companies matched	110	2003-2008	234,802	165	5%	46%

The control group

Defining the control group has some issues due to the selection bias that appears when performing matching analysis. The perfect impact assessment would require that we were able to observe the exact same companies that participate, but simply without the participation. This is unfortunately impossible; therefore we need to make choices regarding the control group. The following criteria are:

- Participation in the national public-private research system at some point
- The companies are in one of the nace 2 categories, where FP6/FP7 participation has been observed.
- Below 1000 full time equivalent employees
- Minimum one master or PhD educated has to be employed in the company.

The first criterion for selection is that the control group has to have some kind of relation to the national public-private research system. Therefore the companies in the control group must have participated in one of the following research voucher, the Danish National Advanced Technology Foundation, Innovation Consortia, Innovation Voucher Scheme or Open Funds. This is due to the fact that the participation in national public-private research projects indicates to some extent that the companies could also have participated in the European framework instead or as a supplement. However three years up until FP6/FP7 participation no other significant participation influences the FP6/FP7 participation.
The resulting sample consists of participating firms that are matched with highly similar firms not participating that we can observe for five years. Five participating firms have had other collaboration activity within the national framework for public-private partnerships in the past three years leading up to the matching point. We have chosen to include them in the sample because 1) they cannot disturb the overall picture, and 2) participation is nationally oriented and of smaller magnitude than their FP6/FP6 activity. Only 2 firms participate more than once in the matching sample, but for both firms, the participation is five years apart (participation in 2003 and again in 2008).

Industries: Nace categories

The impact assessment only includes industries where FP6/FP7 participation occurs. Therefore some of the Nace 2 categories are disregarded. The following table illustrates, which Nace 2 categories are included in the analysis and which sectors these belong to.

Table 0.2
Construction of included industries

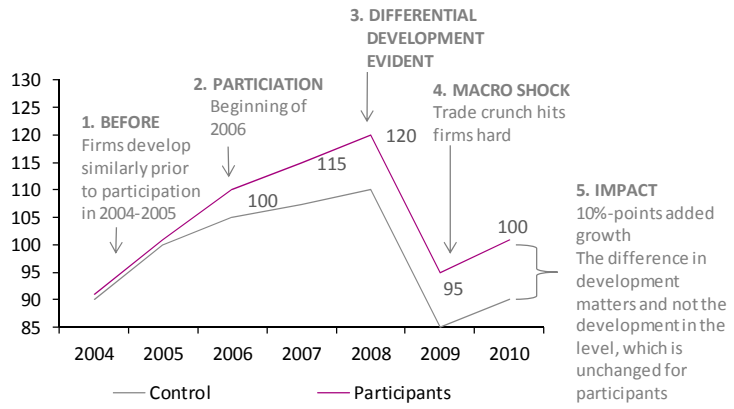
Industry dummies	Sector	Nace 2
C1	Manufacturing	10, 11
C10	Manufacturing	26, 27
C11	Manufacturing	28, 29, 30
C13	Manufacturing	32
C3	Manufacturing	13, 14, 15
C4	Manufacturing	16, 17
C6	Manufacturing	19, 20
C7	Manufacturing	21
C8	Manufacturing	22, 23
C9	Manufacturing	24, 25, 26
F1	Construction	41, 42, 43
G1	Wholesale and retail	45, 46, 47
J1	Information og kommunikation	58
J4	Information og communication	62
M2	Professional, scientific and technical activities	70
M3	Professional, scientific and technical activities	71

Propensity score matching

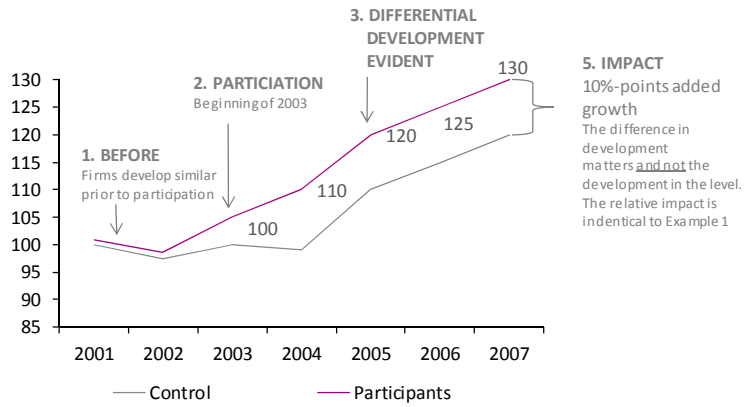
The propensity score matching is performed with common support and replacement. A matching analysis is a counterfactual analysis. For each participating firm we find comparable firms that do not participate (i.e. the control group of firms) but might as well have participated. These non-participating firms constitute a counterfactual scenario for the participating firms, i.e. what we could expect would have happened, had they not participated in FP6/FP7-projects. A difference-in-difference matching analysis estimates whether pairs of participants and non-participants develop differentially and not whether the level of performance variable eventually changes. This is important because underlying trends may influence, especially when we evaluate firms across time and thus different economic cycles.

Figure 1 illustrates one important event in our evaluation period: the international trade crunch. The sampled firms were highly dependent on exports and were probably likely to suffer from the international trade crunch in 2009. The first example illustrates how we might observe that a firm performance had not changed for participating firms over five years, yet we still have an impact because control firms were hit similarly by the trade crunch. The difference-in-difference estimation method takes out this common trend. Thus, even though a performance variable shows zero growth, the differential development relative to control firms may be substantial.

Example 1: Time window for firms influenced by the treade crunch in 2009



Example 2: Time window for firms before the trade crunch in 2009



The probit estimation constructs the pscores, which are used in the final matching analysis. We match 1-1 on industry, year, export status, employment categories to the extent possible. A few of the matches are across the employee intervals, but are evaluated to be fair matches. One exam-

ple is a match between a company with 400 employees and a 600 employee company. Some of the matches are manually made, to make sure to achieve good matches.