Paper to be presented at DRUID17
NYU Stern School of Business, New York, June 12-14, 2017

subsidized and non-subsidized R&D projects: do they differ

Mila Koehler
ZEW Centre for European Economic Research
Innovation economics and industrial dynamics
Mila.koehler@zew.de

Bettina Peters
ZEW Centre for European Economic Research
Innovation Economics and Industrial Dynamics
b.peters@zew.de

Abstract
Little is known about whether and to what extent the outcome of subsidized and non-subsidized R&D projects differ. In this paper we exploit a novel dataset of patent applications filed in Germany between 1995-2005, which allows us to identify if a patent stems from a subsidized project or not. We use a variety of patent indicators to elucidate to what extent successful subsidized and non-subsidized R&D projects differ. Results show that patent applications from subsidized projects have a higher private value, are more often co-applied, more general but less original, and have larger inventor teams when compared to all other patent applications filed by the same firms. These differences seem to reflect that thematic R&D programs aim to support R&D projects that have an immediate economic utilization of results and in which firms collaboratively develop basic technologies in high-tech fields.

Jelcodes:O31,H25
Subsidized and non-subsidized R&D projects: Do they differ?

January 2017

Abstract

Little is known about whether and to what extent the outcome of subsidized and non-subsidized R&D projects differ. In this paper we exploit a novel dataset of patent applications filed in Germany between 1995-2005, which allows us to identify if a patent application stems from a subsidized project or not. We use a variety of patent indicators to elucidate to what extent successful subsidized and non-subsidized R&D projects differ. Results show that patent applications from subsidized R&D projects have a higher private value, are more often co-applied, more general, but less original, and have larger inventor teams when compared to all other patent applications filed by the same firms. These differences seem to reflect that thematic R&D programs aim to support R&D projects that have an immediate economic utilization of results and in which firms collaboratively develop basic technologies in high-tech fields.

JEL: H25, O31, O34, O38, H50
Keywords: R&D, R&D projects, innovation, subsidies, public policy evaluation, patents
1. Introduction

Direct R&D subsidies are the main policy measure to support private R&D investments in Germany (BMBF 2012) as there are e.g. no R&D tax credits. Public support by the federal government to industrial R&D ranged between 1-1.3 billion euro per year in the period 1995-2005 which corresponded to 4%-10% of total annual industrial R&D expenditures (BMBF 2010). Such massive public support raises not only the question whether or not private R&D investments are crowded out but also how subsidized R&D projects differ from purely privately financed ones. Economists usually evaluate the impact of R&D subsidies on firms’ R&D inputs and outputs. The majority of empirical studies find that publicly supported firms exhibit higher R&D inputs compared to non-supported firms (see e.g. the survey in Zúñiga-Vicente et al. 2014). Subsidized firms have also higher R&D outputs such as the number of patents filed, while the productivity of R&D in terms of output per euro of publicly and privately spend R&D is the same (see e.g. Czarnitzki and Hussinger 2004).

Studies using firm-level data, however, cannot answer the question which types of R&D projects receive funding within firms and whether they differ from non-subsidized projects. In principle, differences between subsidized and non-subsidized R&D projects may result from the following reasons: First, firms select with which projects to apply for public funding. Second, the funding agency selects which projects it finally supports. Third, the subsidy itself may change the character of the R&D project and affect its results. Do firms for instance choose to apply with projects that have a low expected private value but high social value or do they seek funding for projects with a higher expected private value? Does the funding agency select projects with larger knowledge spillovers to society or does it rather pick the winners, i.e. those R&D projects with high expected private value to the subsidized firms? And do subsidies enhance the quality of the research project because the firm can afford to employ better researchers, instruments or equipment or because the funding agency gives the subsidized firm access to better networks and information?

So far not much is known about differences between R&D projects that are subsidized and those that are not. This is mostly due to the unavailability of project-level data. Our paper aims at filling this gap and benefits from a unique dataset containing information on the outcome of R&D projects in terms of 174,311 patent applications from subsidized and non-subsidized projects filed at the German Patent Office (GPO) in the period 1995-2005. This data set allows us to shed light on the question whether and to what extent subsidized and non-subsidized R&D projects differ.

Accounting for firm fixed effects, our results show substantial differences in various patent characteristics between subsidized and non-subsidized projects: patents from subsidized R&D projects are on average more general, but less original, and exhibit a higher private value. In addition, they have more often co-applicants and a larger number of inventors which both indicate a higher degree of collaboration. We argue, that this finding reflects the orientation of German federal thematic R&D programs1 to support collaborative R&D projects with a likely economic utilization of the innovation

---

1 Thematic R&D programs are targeted to support R&D projects in high-tech fields such as biotechnology, ecological and sustainable development, health and medicine, aerospace research, information technology, and large equipment for basic research. In some papers such as Czarnitzki and Lopes-Bento (2014), they are referred to as “direct subsidies” which literally translates the German term “Direkte Projektförderung (DPF)”.
aiming to develop basic technologies in promising high-tech fields. Interestingly, except for generality these differences survive if we limit the control sample of patent applications to the same technology classes and application period as patents from subsidized projects of the same firm. Thus, even within the same firm and technological field, subsidized and non-subsidized projects that result in patent applications do differ. This finding is in contrast with results of output additionality studies conducted in a similar time period in Germany which demonstrate that publicly induced R&D expenditure shows a positive but similar productivity as privately induced R&D in terms of the likelihood to file a patent and sales share with market novelties (Aschhoff 2009, Czarnitzki and Hussinger 2004, Czarnitzki and Licht 2006).

This paper is structured as follows. Section 2 briefly explains how R&D subsidies are distributed in Germany. Section 3 reviews the relevant literature and highlights how this study contributes to it while section 4 describes our empirical approach. The data and variables are explained in section 5. Descriptive and estimation results are presented in section 6. Section 7 concludes.

2. Public support for R&D in Germany

In this section we briefly describe why governments distribute R&D subsidies to support private R&D projects and explain how R&D subsidies are distributed in Germany. For our research question, Germany is an especially interesting case to study, as – in the absence of R&D tax credits – R&D subsidies are the main policy tool to support industrial R&D in Germany (BMBF 2012). Given the crucial role of R&D for competitiveness and growth, the German government – like most other governments worldwide – supports private R&D expenditures with public money because there is an assumed undersupply of R&D activities by the industry from a social point of view. Economic theory stresses that firms have insufficient incentives to invest into R&D because they have to bear the full costs while they cannot fully appropriate the returns of their R&D efforts (Arrow 1962). Newly created knowledge may for example leak out of the firm even before a new product is launched to the market, or a product that is successful on the market may be easily imitated by competitors who did not bear the high invention costs which have to be paid by the pioneering firm (Mansfield et al. 1981, Mansfield 1985). Knowledge spillovers to other firms also occur in form of cumulative inventions which build upon the knowledge of the original invention. Besides, R&D investments are to a large extent investments into personnel. As a result, firms bear the risk that with researchers changing jobs tacit knowledge leaves the firm. Hence, firms may be reluctant to invest into R&D as knowledge can easily spill over to rivals and resulting inventions cannot always be fully protected for example by intellectual property, secrecy or complex product design. In addition to spillovers, firms may be financially constrained because of information asymmetries between the firm and outside financiers, high uncertainty of expected returns to R&D, and the fact that firms may lack collateral because R&D investments are mostly sunk as the largest fraction is wages for R&D personnel.

R&D subsidies to the industry are distributed either via thematic or generic R&D programs (BMBF 2012). Thematic R&D programs, so called direct project funding (Direkte Projektförderung), are aimed at stimulating R&D projects in technological areas that the federal government views as particularly important for the future technological competitiveness of Germany. Currently, examples include biotechnology, information technology, aerospace, ecological and sustainable development, health and
medicine, and large equipment for basic research. In contrast, generic R&D programs, so called indirect project funding (Indirekte Projektförderung), are not directed to specific technological areas. These programs are especially targeted on SMEs and firms located in eastern Germany.

The federal ministries – mainly the Federal Ministry of Education and Research (BMBF) and the Federal Ministry for Economic Affairs and Energy (BMWi) – design the R&D programs. In the granting process they are supported by currently 14 different program administrating agencies, which have the scientific and technical expertise to administer and guide the R&D programs (BMBF 2010). After firms have applied for funding, the government decides with the support of the scientific experts at the administrating agencies which R&D project(s) to support. General selection criteria are: compliance with the thematic outline of the R&D program; relevance of the research goal for society; high quality of the R&D project; sound know-how and expertise of the applicant (and its partners); economic and technical potential as well as risk of the project; a likely economic utilization of the innovation and collaboration between different research partners (Fier 2002, Aschhoff et al. 2012). Besides, the firm has to prove its ability to pay its own share of the project costs (Aschhoff et al. 2012). To which extent the state may subsidize industrial R&D projects is strictly regulated in the Community Framework for State Aid for R&D (EC 2006). It foresees that the subsidy amount may at most cover 50% of the total project costs. A markup may be paid for SMEs, collaborative projects and firms located in Eastern Germany. Given the framework of subsidy allocation to private R&D in Germany, it may be that firms strategically apply for funding with collaborative projects of great importance for society in specific technological fields.

3. Literature review and contribution

There are only a few existing studies dedicated to analysing differences between subsidized and non-subsidized R&D projects. This is mainly due to the fact that data on project level is usually unavailable. The only studies are Schneider (2008) and Takalo et al. (2013a). Schneider (2008) analyses differences in the outcome of subsidized and non-subsidized R&D projects. He compares patents from subsidized and non-subsidized projects filed by Danish inventors using data from the PatVal² survey, in which inventors were asked whether or not the patent resulted from a subsidized project (Giuri and Mariani 2005). For a sample of 495 Danish patents, he finds that patents from subsidized R&D projects are more important in terms of forward citations than patents from non-subsidized projects. One limitation of this study is that unobserved firm heterogeneity is not accounted for since patents in the control group are from different firms than patents from subsidized R&D projects. Takalo et al. (2013a) analyse the funding decision of the granting agency in Finland. For a sample of R&D projects with which firms applied for R&D funding, they find that projects of successful applicants are on average rated higher regarding their expected technological challenge than projects of rejected applicants, while rejected and accepted projects are rated similar in terms of their expected economic risk. The latter effect, however, is only weakly significant. Takalo et al. (2013a) only make use of project-level information of submitted R&D projects while they do not observe R&D projects firms did not seek public support for. Thus, nothing is

² The PatVal database comprises information on a sample of European inventors holding a patent at the European Patent Office (EPO).
known so far about how subsidized R&D projects differ from non-subsidized R&D projects conducted by the same firm.

Our work is also related to the vast literature on R&D subsidies, which usually analyses the effect of R&D subsidies at the firm-level not the project-level. One major concern with public R&D subsidies is that public money crowds out private R&D investments, i.e. private R&D expenditure is substituted with public money. Therefore many studies evaluate the effect of R&D subsidies on R&D input by trying to answer an essentially counterfactual question: How much would a subsidized firm have invested in R&D in absence of the subsidy? The major challenge evaluation studies are confronted with is to account for the selection bias in order to identify the causal effect of the subsidy. A selection bias occurs if in particular highly innovative firms apply for a subsidy or if the funding agency chooses to support the firms with the most promising R&D projects out of the pool of subsidy applicants following a “picking-the-winner” strategy.3 Studies that analyse differences between subsidized and non-subsidized firms indeed find that subsidized firms are more successful than non-subsidized firms. In Germany, for instance, firms are more likely to receive R&D funding if they have better trained personnel, compete on international markets but are not foreign owned, conduct R&D continuously, have more innovative capacities and already received public funding in the past (see e.g. Aschhoff 2010, Hussinger 2008, and Czarnitzki and Lopes-Bento 2014).

Input additionality studies compare the R&D expenditure of a subsidized firm to the R&D expenditure in the counterfactual situation and usually find that subsidized firms spend more on R&D than non-subsidized firms (for an overview see the surveys of Zúñiga-Vicente et al. 2014 or David et al. 2000; and for studies on Germany see Czarnitzki and Fier 2002, Almus and Czarnitzki 2003, Aerts and Schmidt 2006, Czarnitzki et al. 2007, Hussinger 2008, Aschhoff 2009, Czarnitzki and Lopes-Bento 2014, Hud and Hussinger 2015).4 Some studies tested if there is additionality, i.e. if subsidized firms spend not only more on R&D than in the counterfactual situation (test full-crowding out) but spend more on R&D after subtracting the amount of the received subsidy (test additionality). Czarnitzki and Hussinger (2004), Aschhoff (2009) and Hud and Hussinger (2015) find on average additionality of public R&D subsidies. However, one has to keep in mind that input additionality is not a necessary condition for a welfare-improving R&D subsidy.

It is unclear though whether R&D subsidies should increase the innovative output of a firm. Part of the R&D subsidy may flow into increased wages of R&D personnel without increasing R&D output productivity since the demand for R&D personnel is rather inelastic in the short run (Goolsbee 1998). In addition, subsidized R&D projects may exhibit a lower expected rate of return, since projects with a high expected rate of return may be more likely privately financed.5 On the contrary, funding agencies may

---

3 Subsidy evaluation studies account for selection employing a variety of different econometric techniques like instrumental variables, matching and selection models.
4 Only one study for Germany (Czarnitzki et al. 2007) finds full crowding out of private investments for a small subgroup of firms that receive a subsidy but do not collaborate at the same time. For the group of collaborators – which is the largest group among subsidized firms – they find higher R&D spending compared to firms that receive no subsidy.
5 Note that even R&D projects with high expected returns may not be conducted due to financial constraints. Firms may lack the means to cover entry costs to start engaging in R&D like building and equipping research facilities and
follow a “picking-the-winner” strategy by choosing highly promising technology areas for thematic R&D programs and by subsidizing R&D projects with a high expected return that result more likely in patentable inventions. Using Japanese data, Branstetter and Sakakibara (1998, 2002) find that participation in a publicly funded R&D consortium increases the participants’ patenting activity. For Germany, there is evidence that the productivity of R&D is similar for publicly and privately financed R&D. This has been shown for the probability to file a patent per euro of R&D (Czarnitzki and Hussinger 2004, Czarnitzki and Licht 2006) and the sales share with market novelties (Aschhoff 2009). In a recent study, Czarnitzki and Lopes-Bento (2014) examine the effect of a subsidy on the quality of patents. Using German data, they find that patents of firms benefitting from national R&D support receive on average more forward citations. Even though they conclude that “public funding triggers socially beneficial research projects” (p.380), it is not clear if this finding is driven by patents from subsidized or non-subsidized R&D projects since firm-level studies fail to distinguish between output from subsidized and non-subsidized R&D projects. However, to disentangle the effect of publicly and privately induced R&D, project-level data is required.

We contribute to the literature by shedding light on the question of whether and to what extent subsidized and non-subsidized R&D projects that result into a patent application differ. Such differences may arise for three reasons. First, firms choose R&D projects for which they seek public funding. However, it is not clear which strategy a firm pursues for project selection in order to maximize the likelihood of getting funding. On the one hand, a firm might choose to apply with a highly profitable R&D project. On the other hand, it might refrain from choosing a project that has a high private value in order to prevent knowledge leakage during the application process. Firms may also decide to apply with a project that is difficult to finance because it is very risky and the created knowledge is difficult to appropriate so that it might generate higher social value. Second, the funding agency selects the R&D projects according to certain criteria. In Germany, R&D funding is targeted to high quality projects with a high value for society, as well as to collaborative projects. It is unclear though if agencies indeed pick projects with a high value for society or tend to fund projects that are rather privately profitable. The project selection behaviour of the agency will in turn affect a firm’s choice of projects to apply for subsidies. Third, there may be an effect of the subsidy on the project in a sense that receiving a subsidy enhances the project outcome. The subsidy may for instance enable a firm to employ better researchers, instruments or equipment resulting in a higher quality of the research project. The project outcome might also be affected if the funding agency gives the subsidized firm access to better networks and information.

We investigate whether subsidized and non-subsidized R&D projects that result in patent applications differ in several key characteristics such as their private value, knowledge spillovers, basicness and project size. We use a comprehensive sample of 174,311 patent applications filed in Germany between 1995 and 2005 by 249 firms, which have filed at least one patent application from a subsidized R&D project. We assume that ex post realized inventions in form of patent applications allow us to learn about the characteristics of the underlying R&D projects. Our specific sample design allows us to control hiring and training researchers. Arqué-Castells and Mohnen (2012) find e.g. that 10% among non-R&D performing manufacturing firms in Spain might become R&D performers in the long run, if entry costs were covered by public authorities.
for unobserved firm heterogeneity, because we compare patent applications from subsidized and non-
subsidized R&D projects of the same firms. As all firms in our sample are subsidy receivers, we do not
have to worry about selection bias arising from a firm’s selection into receiving a subsidy.

4. Empirical approach

In order to analyse the characteristics of R&D projects for which firms receive public funding, we
estimate a Probit model on the probability that a patented invention stems from a subsidized R&D
project. We expect this decision to mainly depend on a patent’s private value, knowledge spillovers,
basicness and the size of the project. In section 5.3 we describe in detail how we measure these
characteristics. The outcome whether or not a patent stems from a subsidized project is at first place
driven by a firm’s choice which projects to select for application. It is also affected by the agency’s
decision to subsidize the project but we assume that their decision depends on the same factors (see
section 2). As we cannot rule out reverse causality, we do not intend to interpret our estimation in a
causal way. Instead we conduct Probit estimations as a multivariate descriptive analysis on differences
between patents from subsidized and non-subsidized R&D projects. This allows us to learn about the
differences of successful subsidized and non-subsidized projects. The estimated Probit model can be
described as

\[(1) \quad P(\text{SUBS}_i) = \Phi(X'_i\beta + Y'_i\beta + Z'_i\beta),\]

where \(\text{SUBS}_i\) takes value one if patent \(i\) stems from a subsidized R&D project and zero otherwise. \(X'_i\) is a
vector of patent characteristics that captures the expected private value, knowledge spillovers,
basicness, and the size of a R&D project. \(Y'_i\) is a vector of year and technological area dummies, and \(Z'_i\) is
a vector of firm dummies. By including year and technological area dummies, we control for differences
in the probability to file a patent from a subsidized R&D project over time and technology fields. This is
important because not only the government’s budget for R&D subsidies to the industry changes over
time but also to which extent certain technological areas are targeted by federal R&D programs.

The probability of receiving R&D funding and hence of filing a patent application from a subsidized R&D
project varies across firms. If firm-specific characteristics are unobserved but correlated with the
explanatory variables, this will lead to biased results. For instance, larger firms are more likely to file
triadic patents, i.e. patents filed in the EU, Japan and US, and have at the same time a higher probability
of receiving public R&D funding due to their experience in innovation and in attracting public money.
We control for this unobserved firm-specific heterogeneity by including firm dummies into the
regression model. This way we investigate if correlations between patent indicators and the propensity
to file a patent from a subsidized R&D project are spurious or robust to controlling for firm-specific
effects.

We will estimate the model for two different samples which are explained in more detail in section 5.1.
Sample 1 consists of all patent applications by firms that have filed at least one patent from a subsidized
R&D project. This allows us to investigate differences between patent applications from subsidized R&D
projects and the rest of a firm’s patent portfolio. Sample 2 contains only patent applications that stem
from the same technological area and application period as the patent application(s) from subsidized
projects of the same firm. Thus, we can analyse whether or not differences between patent applications
from subsidized and non-subsidized R&D projects prevail if filed in the same technological area. As thematic R&D programs in Germany are targeted to specific technological areas, the comparison of patents in sample 2 allows us to partially disentangle the selection effect from the effect a subsidy might have on project outcomes. If we find differences between subsidized and non-subsidized projects in sample 1 which disappear in sample 2, the differences are likely to reflect a selection effect. Firms select projects from technological areas that comply with the orientation of thematic R&D programs. On the other hand, if differences are more pronounced in sample 2, this may point towards an impact of the subsidy on project characteristics.

5. Data and variables

5.1 Construction of the patent database

The following empirical analysis is based on patent application data. We created a unique database that contains information on patent applications at the German Patent Office (GPO) between 1995 and 2005. The crucial and novel information is whether or not a patent application stems from a subsidized project. By collateral clauses firms are obliged to indicate to the GPO that the invention originates from a subsidized project funded by a thematic R&D program when seeking patent protection at the GPO. That is, when filing a patent at the GPO firms have to hand in the special form “Information on the Funding Project Number for Intellectual Property Applications” that indicates the source of funding (federal ministry), the project name and the official project number. It has only been recently that the GPO revealed the information if a patent application resulted from a subsidized project in their online database. We use this information as a starting point for constructing a unique database.

As a first step we identified 1,566 patent applications that stem from subsidized R&D projects in the GPO database between 1995 and 2005. We will denote them as subsidized patents hereafter. Second, we manually identified the patent applicant by name search as patent applicants are not given a unique applicant number by the patent office. That is, we ran a computerized name search that bundled all similar applicant names and manually selected which applicant names belong to the same applicant. Third, we manually identified if the applicant is a firm, private person, university or research institute. We only keep firms in our final sample since we are interested in studying differences between subsidized and non-subsidized R&D projects conducted by the private sector. This resulted in a sample of 1,230 subsidized patent applications filed by 271 different firms.

As a next step, we drew our group of control patents. Control patents are all non-subsidized patents filed by the same 271 firms between 1995 and 2005 at the GPO. This way we can analyse differences between subsidized and non-subsidized R&D projects within subsidized firms. We complement the dataset by adding further information for both subsidized and non-subsidized patents from PATSTAT such as the priority application date, IPC-classes, citations made to other patents (backward citations)

6 The collateral clauses are a general part of receiving cost-based grants and they specify auxiliary terms and conditions for funds provided by the Federal Ministry of Education and Research to commercial companies for R&D projects. Between 1988 and 1997 the collateral clauses were called NKFT88. In 1998 a reform of the collateral clauses were implemented, the so called NKBFP98.
and citations received by other patents (forward citations), co-applicants (y/n) and countries patent protection is applied for. We dropped double counts of collaborative patent applications. They occur if a collaborative patent was filed by two or more of the firms in our database. Moreover, we excluded 23 subsidized patents filed by 22 firms from our sample because these firms did not file any control patent between 1995 and 2005. The final dataset contains a total of 174,311 patent applications filed by 249 firms between 1995 and 2005; thereof 1,207 applications originate from a subsidized R&D project (0.7%) and 173,104 from a non-subsidized R&D project. We call this dataset sample 1.

Patent filing behaviour as well as citation patterns are likely to change over time and technology fields. Furthermore, thematic R&D programs in Germany are targeted to support innovation projects in specific technologies. As already explained in the previous section, we therefore created a subsample (sample 2) which further restricts the control group to patent applications filed by the same firm within the same technological area and the same application period as the subsidized patent application. For that purpose, we assigned the IPC classes to 30 different technological areas according to Schmoch (2008). As application period we define a five-year window, i.e. up to two years before or after the application year of the subsidized patent. This rule was chosen because there are not always enough control patents filed in the same application year. For 49 subsidized patents there are no control patents within the same technological area and application period so we dropped these observations from sample 2. In total, sample 2 contains 89,905 patent applications filed by 218 firms of which 1,158 stem from subsidized R&D projects (1.3%).

5.2 Sample characteristics

Firms in our sample differ strongly in their patenting behaviour. While some firms filed only two patent applications within the 11 years of observation, other firms filed a few thousand. Table 1 displays the number of patent applications filed per firm in sample 1 in the period 1995-2005. The first column shows the distribution of patent applications from subsidized projects across firms while the second column depicts the distribution of patent applications from non-subsidized projects. Table 1 demonstrates that the amount of applications filed per firm is highly skewed. Half of the firms, for instance, filed only one patent application from a subsidized project, 75% of the firms have up to 3 applications, while 5% of firms filed 12 applications and more. An even more skewed pattern emerges among our control group of patent applications from non-subsidized projects. 10% of the firms filed up to 3 control patents while the most active 10% of the firms filed 1,229 and more patent applications from non-subsidized R&D projects. We find similar results for sample 2.

Column 3 of Table 1 highlights that the share of patents from subsidized R&D projects in all patents filed by the same firm is likewise heavily skewed. For 5% of the firms the share of patents from subsidized R&D projects in their patent portfolio is 0.11% or below. This implies that there are 900 or more control patents for each subsidized patent. In contrast, for the top 5% of the firms we observe a share of patents from subsidized R&D projects of at least 50%, i.e. there are two or less control patents for each subsidized patent. The skewed distribution of patent applications is likely to be highly correlated with firm size. To account for the skewed distribution of patent applications in our sample we will carry out robustness checks for low patenting firms and high patenting firms (see section 6.3).
Table 1: Distribution of the number of patent applications per firm and the share of patents from subsidized R&D projects per firm between 1995 and 2005 (sample 1)

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Number of patent applications per firm from...</th>
<th>Share of patents from subsidized R&amp;D projects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>...subsidized R&amp;D projects</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>1</td>
<td>0.01%</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; Percentile</td>
<td>1</td>
<td>0.05%</td>
</tr>
<tr>
<td>5&lt;sup&gt;th&lt;/sup&gt; Percentile</td>
<td>1</td>
<td>0.11%</td>
</tr>
<tr>
<td>10&lt;sup&gt;th&lt;/sup&gt; Percentile</td>
<td>1</td>
<td>0.25%</td>
</tr>
<tr>
<td>25&lt;sup&gt;th&lt;/sup&gt; Percentile</td>
<td>1</td>
<td>1.09%</td>
</tr>
<tr>
<td>50&lt;sup&gt;th&lt;/sup&gt; Percentile</td>
<td>1</td>
<td>4.77%</td>
</tr>
<tr>
<td>75&lt;sup&gt;th&lt;/sup&gt; Percentile</td>
<td>3</td>
<td>16.67%</td>
</tr>
<tr>
<td>90&lt;sup&gt;th&lt;/sup&gt; Percentile</td>
<td>6</td>
<td>33.33%</td>
</tr>
<tr>
<td>95&lt;sup&gt;th&lt;/sup&gt; Percentile</td>
<td>12</td>
<td>50.00%</td>
</tr>
<tr>
<td>99&lt;sup&gt;th&lt;/sup&gt; Percentile</td>
<td>60</td>
<td>63.64%</td>
</tr>
<tr>
<td>Maximum</td>
<td>287</td>
<td>88.89%</td>
</tr>
<tr>
<td>No. of firms</td>
<td>249</td>
<td>249</td>
</tr>
</tbody>
</table>

Note: All values are calculated using sample 1. The share of patents from subsidized R&D projects is defined as the number of subsidized patents divided by the total number of patents filed by a firm.

5.3 Variables

Dependent variable. Our dependent variable is the dummy variable SUBS that takes unit value if a patent application stems from a subsidized R&D project and zero otherwise.

Patent indicators. We aim at testing whether patent applications from subsidized R&D projects differ from the control patent applications in four key characteristics: private value, knowledge spillovers, basicness and project size. Each of these key characteristics is proxied by a variety of different patent indicators that have been widely used in the literature. Ultimately, politicians are interested in the social value of a project which corresponds to the social value of a project’s outcome measured by patent applications in our data. The social value is composed of the private value and spillovers to the economy not captured in the private value, which we approximate by knowledge spillovers and basicness indicators.

Private value. We measure the private value of a patent application using two different indicators: self-citations and triadic patents. Self-citations (SELFCIT) are citations a patent application received by subsequent patent applications filed by the same firm. They reflect follow-up developments of the original invention by the same firm. These cumulative inventions signal to what extent a firm manages to appropriate returns from the cited invention. Self-citations therefore signal the profitability of the cited invention, because “the higher the proportion of these later developments that take place `in-house´ the larger would be the fraction of the benefits captured by the original inventor” (Trajtenberg et al. 1997, p.29). Hall et al. (2005) demonstrate that self-citations are positively correlated with a firm’s market value measured by Tobin’s q. We measure self-citations by counting only citations from patent
applications filed at the GPO or EPO as both patent offices have high and comparable standards. We eliminate double counts of citations by using only one citation per patent family.\(^7\) Moreover, we restrict the forward citations to five years. The five-year window was chosen because early citations, i.e. those observed within five years of application date, are highly correlated with the importance and economic value of patents, while correlation is much lower for later citations (Lanjouw and Schankerman 1999).

**Triadic patents** (TRIADIC) are patents filed in all three major markets: the US, Europe and Japan. They reflect the quality of the patent (Nagaoka et al. 2010). Given the costs attached to filing a patent, a firm’s willingness to seek patent protection at all three patent offices, reveals a higher valuation for that invention, indicating an overall higher quality or private value of that patent application.

**Knowledge spillover.** Knowledge spillovers are captured by two indicators: external citations and co-applications. **External citations** (EXTCIT) are citations received by subsequent patent applications of another firm. As with self-citations, we measure external citations by counting only citations from patent applications filed at the GPO or EPO, eliminating double counts and restricting the forward citations to five years after the patent application was filed. In line with Jaffe et al. (1993) and Jaffe et al. (2000) we interpret external citations as partial measure of knowledge spillovers to other firms. Admittedly, using forward citations is a noisy measure of knowledge spillovers. On the one hand there may be spillovers without citations since not all knowledge spillovers trigger inventions filed at the GPO or EPO for patent protection. But on the other hand there may be citations without spillovers. At EPO and GPO citations are included by the patent examiner and he may have added references that are related to the application content-wise but did not serve as a source of inspiration to the inventing firm. Despite the obvious flaws, forward citations have the advantage of leaving a measurable “paper trail” of an otherwise unobserved phenomenon (see Jaffe et al. 1993). The key question for drawing inferences about knowledge spillovers from citations in our research is whether we expect that (i) citations without spillovers and (ii) spillovers without citations systematically occur more or less often for patent applications from subsidized than from non-subsidized projects. We cannot think of any good reason why these should differ for patents from subsidized and non-subsidized R&D projects.

Furthermore, knowledge is intendedly exchanged in collaborative projects. **Co-application** (COAPP) indicates collaboration with another firm, university or research institute. However, not all collaborations are reflected in co-applications of the project partners. Co-applications involve legal complexities managing the IP rights across partners, making it costly for firms to co-patent (Hagedoorn 2003). Based on the PatVal survey, inventors state that 15% of the filed patents stem from co-inventing with other organizations, while only 3.6% of all patents resulted in co-applications (Giuri and Mariani 2005). Even though an imperfect measure, co-applications still reflect a significant share of co-inventions and thus knowledge exchange. Again the crucial question for our analysis is whether co-inventions from subsidized R&D projects result more or less often in co-applications than co-inventions from non-subsidized R&D projects. We cannot think of any good reason why this should be the case.

\(^7\) A patent family comprises patents that cover the same invention but are e.g. filed at different authorities or with slight adjustments at different points in time.
Basicness. The basicness of an invention is approximated by three indicators: generality, originality (Trajtenberg et al. 1997) and the number of backward citations. Generality (GENERAL) is defined by one minus the Hirschman-Herfindahl index of the citing 3-digit technology classes:

\[
\text{Generality}_i = 1 - \sum_{k=1}^{N_i} \left( \frac{\text{Citing}_{ik}}{\text{Citing}_i} \right)^2
\]

Where **Citing** are the number of citations within 5 years to patent \(i\), \(k\) indicates the 3-digit technology classes and \(N_i\) the number of different 3-digit technology classes to which the citing patents belong. Generality is typically interpreted as a measure of a technology’s basicness. A technology is assumed to be more basic, the more inventions it triggers in a broader range of technological fields. The generality index ranges between zero and one with higher values representing more generality or basicness implying that the invention underlying the original patent is of wider technological applicability. We also create a dummy variable NOFORW that takes on the value one if a patent received no forward citation, because the generality index is set to zero in that case.

In contrast to generality, the originality of a patent application is captured by the concentration of citations to predecessors. Thus, originality (ORIGINAL) is defined as one minus the Hirschman-Herfindahl index of the cited 3-digit technology classes:

\[
\text{Originality}_i = 1 - \sum_{k=1}^{N_i} \left( \frac{\text{Cited}_{ik}}{\text{Cited}_i} \right)^2
\]

Where **Cited** are the number of backward citations to prior art by the patent \(i\) in the 3-digit technology class \(k\). An invention is assumed to be more original or complex, if it draws from many different technological areas. The originality index ranges between zero and one with higher values representing a wider variety of cited technological fields. We also generate a dummy variable NOBACK that equals one if a patent cites no prior art, because the originality index is set to zero in that case.

Backward citations (BACKCIT) reflect to which extent the invention draws from pre-existing knowledge. Backward citations are seen as a measure of a patent’s basicness (Trajtenberg et al. 1997). Fewer backward citations indicate that a patent is more basic because it does not cite many predecessors and can thus be assumed to be more creative and original.

Project size. Firms might also seek funding for projects that are larger and hence more difficult to finance privately. We use the inventor team size as a proxy of project size. Inventor count (INVCNT) is the number of inventors listed on the patent application.

Control variables. As citation patterns may change over time and technology classes, we employ time dummies for each year in our sample (YEAR1995-YEAR2005) as well as dummies for 30 different technology classes (AREA1-AREA30). Following Schmoch (2008), we assigned the indicated IPC classes of the applications to 30 different technological areas (see Table 5 in Appendix 1). Note that a patent application can be assigned to more than one IPC class. Applications belonging to more than one technological area are assigned to the area of the majority of its 4-digit IPC subclasses. In case the same amount of IPC subclasses were assigned to two or more technological areas we assign the application to two or more technological areas.
6. Empirical Evidence

This section presents empirical evidence on the question whether and to what extent subsidized and non-subsidized R&D projects differ.

6.1 Descriptive statistics

Based on sample 1, Table 2 displays the summary statistics of patent indicators for all patent applications and for the respective subgroups of subsidized and non-subsidized patent applications. The table also contains a t-test on the mean difference between subsidized and non-subsidized patent applications. We focus on sample 1 because results for sample 2 are very similar.

We find no significant differences with respect to private value indicators of subsidized and non-subsidized patent applications. The average self-citations are slightly higher for applications from subsidized when compared to non-subsidized R&D projects, though the difference is not statistically significant at the 5%-level. Approximately 31% of the applications are triadic patents (TRIADIC) that are filed not only at the EPO but also at the USPTO and the JPO. As in the case of self-citations, we do not find significant differences between subsidized and non-subsidized patent applications.

The evidence for the indicators capturing knowledge spillovers is mixed. Applications receive on average 0.7 external citations (EXCTCIT). Most interestingly and in contrast to Schneider (2008), external citations are significantly lower for subsidized patents than for non-subsidized patents. This finding points towards larger knowledge spillovers of patented inventions from non-subsidized projects. This is surprising as R&D subsidies principally should aim to support projects that exert large external knowledge spillovers and thus are beneficial for society. Our finding may either reflect that our measure of external knowledge spillovers is too noisy, that firms do not apply with ideas that exert higher knowledge spillovers compared to other R&D projects of the same firm, or that the funding agency does not choose projects with high knowledge spillovers (but rather high private value or high spillovers realized through other channels than knowledge spillovers).

Furthermore, we find a significantly higher share of co-applied (COAPP) patent applications (9%) in case of subsidized patents compared to applications from non-subsidized projects (3%) indicating a higher degree of knowledge exchange. This difference is statistically significant on the 0.1%-level and probably reflects the strong focus of the German federal government to support collaborative R&D projects in order to foster the internalization of knowledge spillovers among research partners.

We find mixed evidence also with respect to the variables measuring the basicness of an invention. Regarding generality (GENERAL), applications score on average low with 5%. Furthermore, there is no significant difference between subsidized and non-subsidized patents. Likewise, we find no significant differences in the likelihood to be cited at all as measured by the dummy variable NOFORW. On average 57% of subsidized patents are not cited at all within the first five years after application and this proportion is only marginally smaller for non-subsidized patents (56%).
Table 2: Descriptive statistics of patent applications from non-subsidized and subsidized R&D projects 1995-2005 (sample 1)

<table>
<thead>
<tr>
<th>Variable</th>
<th>All mean</th>
<th>s.d.</th>
<th>min</th>
<th>max</th>
<th>Subsidized mean</th>
<th>s.d.</th>
<th>min</th>
<th>max</th>
<th>t-test</th>
<th>Non-Subsidized mean</th>
<th>s.d.</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELFCIT</td>
<td>0.306</td>
<td>0.994</td>
<td>0</td>
<td>47</td>
<td>0.358</td>
<td>0.972</td>
<td>0</td>
<td>9</td>
<td>0.306</td>
<td>0.995</td>
<td>0</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>TRIADIC</td>
<td>0.307</td>
<td>0.461</td>
<td>0</td>
<td>1</td>
<td>0.294</td>
<td>0.456</td>
<td>0</td>
<td>1</td>
<td>0.307</td>
<td>0.461</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Spillovers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXTCIT</td>
<td>0.728</td>
<td>1.583</td>
<td>0</td>
<td>87</td>
<td>0.593</td>
<td>1.198</td>
<td>0</td>
<td>13</td>
<td>0.729</td>
<td>1.585</td>
<td>0</td>
<td>87 **</td>
<td></td>
</tr>
<tr>
<td>COAPP</td>
<td>0.032</td>
<td>0.177</td>
<td>0</td>
<td>1</td>
<td>0.085</td>
<td>0.278</td>
<td>0</td>
<td>1</td>
<td>0.032</td>
<td>0.176</td>
<td>0</td>
<td>1 ***</td>
<td></td>
</tr>
<tr>
<td>Basicness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENERAL</td>
<td>0.048</td>
<td>0.140</td>
<td>0</td>
<td>0.89</td>
<td>0.053</td>
<td>0.150</td>
<td>0</td>
<td>0.78</td>
<td>0.048</td>
<td>0.140</td>
<td>0</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>NOFORW</td>
<td>0.560</td>
<td>0.496</td>
<td>0</td>
<td>1</td>
<td>0.572</td>
<td>0.495</td>
<td>0</td>
<td>1</td>
<td>0.560</td>
<td>0.496</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>ORIGINAL</td>
<td>0.112</td>
<td>0.210</td>
<td>0</td>
<td>0.96</td>
<td>0.149</td>
<td>0.239</td>
<td>0</td>
<td>0.88</td>
<td>0.112</td>
<td>0.210</td>
<td>0</td>
<td>0.96 ***</td>
<td></td>
</tr>
<tr>
<td>BACKCIT</td>
<td>1.862</td>
<td>2.751</td>
<td>0</td>
<td>44</td>
<td>2.381</td>
<td>2.972</td>
<td>0</td>
<td>28</td>
<td>1.858</td>
<td>2.749</td>
<td>0</td>
<td>44 ***</td>
<td></td>
</tr>
<tr>
<td>NOBACK</td>
<td>0.540</td>
<td>0.498</td>
<td>0</td>
<td>1</td>
<td>0.395</td>
<td>0.489</td>
<td>0</td>
<td>1</td>
<td>0.541</td>
<td>0.498</td>
<td>0</td>
<td>1 ***</td>
<td></td>
</tr>
<tr>
<td>Team size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INVCNT</td>
<td>2.728</td>
<td>1.898</td>
<td>1</td>
<td>26</td>
<td>3.017</td>
<td>1.837</td>
<td>1</td>
<td>14</td>
<td>2.726</td>
<td>1.898</td>
<td>1</td>
<td>26 ***</td>
<td></td>
</tr>
</tbody>
</table>

No. of obs. | 174,311 | 1,207 | 173,104 |
No. of firms | 249 | 249 | 249 |

Notes: All values are calculated using sample 1. Columns with heading s.d. display the standard deviation, min. minimum values and max. maximum values. *p<0.10, **p<0.05, ***p<0.01.

We find that patent applications from subsidized projects cite on average significantly more prior art (2.38) than applications from non-subsidized projects (1.86) (BACKCIT). Likewise, the dummy variable NOBACK shows that on average 54% of the applications from non-subsidized projects do not cite any prior art, while this share is with only 40% significantly lower for patents from subsidized projects. Having more backward citations indicates that a patent application is less original as it builds stronger on prior art. The originality index additionally reveals information about the dispersion of these backward citations across technology areas. Subsidized R&D patents exhibit a significantly higher score on the originality index (15%) than non-subsidized patents (11%). However, it has to be considered that the originality index is set from a missing value to zero if a patent does not cite prior art at all. Hence, we run a t-test on differences in the originality index only for those applications that have at least one backward citation. Given at least one backward citation, we do not detect any significant differences in the originality index of subsidized and non-subsidized patents.

Finally, we find that the inventor team is with on average 3 inventors significantly larger for patent applications from subsidized R&D projects when compared to applications from non-subsidized R&D projects (2.7). It is unclear though if this observation reflects that firms apply with larger projects or if the subsidy increases the project size.

The employed patent indicators are partly based on the same sources. Table 3 therefore depicts the pairwise correlation between the patent indicators for sample 1. The correlation matrix demonstrates that basically all indicators are correlated to each other at the 5%-level. In the majority of cases, however, the correlation is rather low and therefore does not give rise to multicollinearity concerns. Not surprisingly, exceptions are the high correlations between originality and backward citations (0.612) and
generality and external citations (0.496) which may be partially explained by the fact that a patent that received no backward (forward) citations also scores zero on the originality (generality) index.

We therefore calculated the correlation between generality and external citations for patents that receive at least one forward citation (NOFORW=0) and observe a substantially reduced correlation of 0.37. We similarly find a much smaller correlation between originality and backward citations of 0.34 when restricting the sample to applications that cite prior art (NOBACK=0). Hence, we are confident that correlations between the patent outcome indicators do not cause major multicollinearity problems in our estimations once we include the dummies NOFORW and NOBACK.

Table 3: Correlation matrix of patent indicators for sample 1

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SELFCIT</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. TRIADIC</td>
<td>0.087*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. EXTCIT</td>
<td>0.239*</td>
<td>0.037*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. COAPP</td>
<td>0.008*</td>
<td>0.024*</td>
<td>0.026*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. GENERAL</td>
<td>0.267*</td>
<td>0.046*</td>
<td>0.496*</td>
<td>0.021*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. ORIGINAL</td>
<td>0.000</td>
<td>-0.145*</td>
<td>0.059*</td>
<td>0.015*</td>
<td>0.085*</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. BACKCIT</td>
<td>0.020*</td>
<td>-0.182*</td>
<td>0.094*</td>
<td>0.014*</td>
<td>0.058*</td>
<td>0.612*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8. INVCNT</td>
<td>0.050*</td>
<td>0.191*</td>
<td>0.012*</td>
<td>0.125*</td>
<td>0.028*</td>
<td>-0.082*</td>
<td>-0.102*</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: *p<0.05.

6.2 Econometric results

The Probit analysis allows us to learn in which characteristics subsidized and non-subsidized projects differ controlling for firm-, technology- and year effects. We estimate the Probit model on SUBS for sample 1 and 2 (see equation 1) with and without controlling for firm heterogeneity. An important finding is that accounting for firm heterogeneity is essential for the model specification. Including firm dummies substantially increases the model’s goodness of fit as indicated by McFadden’s R² and the count R². All in all, the model seems to fit the data quite well. Table 4 displays the average marginal effects of the Probit estimations.

As discussed before there may be reverse causality at work. Not only may firms strategically choose specific projects to apply with for public funding, but the subsidy itself may change project characteristics if it for example enables firms to employ higher qualified researchers or better technical equipment. We therefore interpret our results as correlations and not in a causal way.

Let us first consider the results for sample 1 (Column 1 and 2 in Table 4). Projects with a higher expected private value - as indicated by a triadic patent application (TRIADIC) - are correlated with a higher likelihood of getting funded by the agency. Recall that a triadic patent application indicates the value of an invention to the firm, as the firm is willing to bear higher application costs compared to an application at merely the GPO. We neither detected this difference in the descriptive statistics nor in the model without firm heterogeneity, which highlights that it is important to account for firm-specific effects. Thus, inventions with a higher expected private value – as indicated by triadic patent applications – have a higher likelihood of being subsidized. However, we do not find a significant correlation between the number of self-citations (SELCFCIT) and the probability of receiving a subsidy.
One interpretation of the latter result is that firms choose projects to apply for public funding based on the expected value without taking future benefits, as measured by self-citations, into account. In a nutshell, we find a higher expected private value to be associated with a higher probability of receiving a subsidy. On the one hand, this finding reflects that firms choose to apply for public funding with projects that have a higher private value as compared to other projects in their portfolio. On the other hand it may indicate that thematic R&D programs are targeted at specific technological areas which are likely to provide higher profit opportunities and thus higher private values of resulting inventions. In order to differentiate between these two hypotheses, we use sample 2 which compares subsidized patent applications to non-subsidized patent applications filed by the same firm within the same technological area and application period (Column 3 and 4, Table 4). This allows us to examine whether differences in the private value persist when comparing technologically similar inventions from subsidized and non-subsidized R&D projects of the same firm. The results for TRIADIC become even larger in sample 2 and remain significant. We can therefore conclude that the higher likelihood of getting funding for projects with higher expected private value is not driven by thematic preferences of agencies but by firms’ decision to apply for public funding with projects that have a higher private value.

Let us next consider the results regarding knowledge spillovers. In theory, public subsidies should be distributed to projects that are of high value for society. Hence, we expect subsidized projects to exhibit larger knowledge spillovers not only through collaboration but also through external knowledge spillovers to follow-up inventions by other firms. However, the econometric evidence is mixed. We find a positive and highly significant correlation between co-applications (COAPP) and the propensity that an R&D project gets publicly funded. This result is robust to including firm dummies and suggests that the agency focusses on supporting knowledge exchange by fostering R&D collaborations. In contrast, we do not find patent applications with high expected knowledge spillovers – as measured by external forward citations (EXTCIT) – to be correlated to a higher likelihood of getting publicly funded. Whereas the descriptive statistics and the regression without firm dummies even pointed to a significantly negative correlation, this effect vanishes when controlling for firm heterogeneity. We find these results to remain valid in sample 2 when we compare technologically similar patents. The question remains open if this finding is due to the noisy measure of external knowledge spillovers, disincentives for firms to apply with more socially valuable projects, disincentives of the agency to select projects with especially high knowledge spillovers, difficulties to judge a project’s knowledge spillovers to society ex ante, and/or that spillovers are realized through unobserved other channels like an increase in consumer surplus.

Regarding a project’s basicness the results of sample 1 show that a higher generality (GENERAL) is significantly correlated with a higher likelihood of receiving a subsidy. This finding is robust to including firm dummies. A higher score on the generality index indicates that follow-up inventions take place in a wider variety of technological areas. The result that patent applications from subsidized projects have a more general applicability reflects the government’s preference to support R&D projects with relevant research goals for society. Once we restrict the control sample to technologically similar patents within firms in sample 2, this effect vanishes. As argued before, the subsidy itself could have a positive impact on a project’s basicness as the funding agency e.g. promotes the dissemination of the newly created knowledge. However, in this case we should observe higher external forward citations, which is not the case. Therefore, our findings are rather affected by firm and agency selection than by the subsidy itself.
Table 4: Average marginal effects of the Probit estimation on SUBS for the period 1995-2005

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample 1</th>
<th>Sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>Private value</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELFCIT</td>
<td>0.00029</td>
<td>0.00016</td>
</tr>
<tr>
<td></td>
<td>(0.00021)</td>
<td>(0.00020)</td>
</tr>
<tr>
<td>TRIADIC</td>
<td>-0.00034</td>
<td>0.00128**</td>
</tr>
<tr>
<td></td>
<td>(0.00046)</td>
<td>(0.00053)</td>
</tr>
<tr>
<td><strong>Spillovers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXTCIT</td>
<td>-0.00038*</td>
<td>-0.00017</td>
</tr>
<tr>
<td></td>
<td>(0.00020)</td>
<td>(0.00018)</td>
</tr>
<tr>
<td>COAPP</td>
<td>0.00869***</td>
<td>0.00413***</td>
</tr>
<tr>
<td></td>
<td>(0.00158)</td>
<td>(0.00142)</td>
</tr>
<tr>
<td><strong>Basicness</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENERAL</td>
<td>0.00326**</td>
<td>0.00345**</td>
</tr>
<tr>
<td></td>
<td>(0.00166)</td>
<td>(0.00166)</td>
</tr>
<tr>
<td>NOFORW</td>
<td>-0.00030</td>
<td>-0.00066</td>
</tr>
<tr>
<td></td>
<td>(0.00051)</td>
<td>(0.00049)</td>
</tr>
<tr>
<td>ORIGINAL</td>
<td>-0.00005</td>
<td>-0.00006</td>
</tr>
<tr>
<td></td>
<td>(0.00110)</td>
<td>(0.00107)</td>
</tr>
<tr>
<td>BACKCIT</td>
<td>-0.00008</td>
<td>0.00002</td>
</tr>
<tr>
<td></td>
<td>(0.00010)</td>
<td>(0.00010)</td>
</tr>
<tr>
<td>NOBACK</td>
<td>-0.00448***</td>
<td>-0.00271***</td>
</tr>
<tr>
<td></td>
<td>(0.00061)</td>
<td>(0.00071)</td>
</tr>
<tr>
<td><strong>Team size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INVCNT</td>
<td>0.00044***</td>
<td>0.00075***</td>
</tr>
<tr>
<td></td>
<td>(0.00011)</td>
<td>(0.00015)</td>
</tr>
<tr>
<td>Year dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Area dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm dummies</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>McFadden's $R^2$</td>
<td>0.079</td>
<td>0.230</td>
</tr>
<tr>
<td>McKelvey &amp; Zavoina's $R^2$</td>
<td>0.145</td>
<td>0.261</td>
</tr>
<tr>
<td>Count $R^2$</td>
<td>64.97%</td>
<td>77.11%</td>
</tr>
<tr>
<td>No. of obs.</td>
<td>174,311</td>
<td>174,311</td>
</tr>
<tr>
<td>No. of firms</td>
<td>249</td>
<td>249</td>
</tr>
</tbody>
</table>

Notes: Standard errors shown in parentheses. *p<0.10, **p<0.05, ***p<0.01. The count $R^2$ describes the share of correctly classified predictions. For its calculation we do not use the usual 0.5 cut-off point but the actual share of subsidized patents in the sample, since we observe only a very small proportion of subsidized patents in the data set.

Like in the descriptive statistics, we find a significant negative correlation between NOBACK and the probability of getting subsidized which is robust to the inclusion of firm dummies. That is, inventions that have at least one backward citation receive more likely public funding. These inventions can be argued to be less basic than inventions that cite no prior art at all. However, there are no differences in the number of backward citations (BACKCIT) and in the dispersion of backward citations (ORIGINAL) between subsidized and non-subsidized patent applications. This holds for both samples 1 and 2.
Hence, regarding basicness indicators our results are twofold. On the one hand we find that patented inventions from subsidized R&D projects are more general affecting future inventions in a wider variety of technologies compared to non-subsidized patents. This result is driven by thematic preferences of agencies. On the other hand we find that non-subsidized patents have more often no backward citations, indicating that subsidized patents are less original compared to non-subsidized patents. The latter result may reflect that the federal government not only aims to support socially valuable projects but that support is also directed towards projects that have a likely economic utilization which will rather be the case for less fundamental research projects. The question is, if such required economic utilization is socially desirable.

Finally, we find a significantly positive correlation between project size – as measured by the number of inventors (INVCNT) – and the likelihood to get publicly funded. This observation is robust to the inclusion of firm dummies and to controlling for the technological area. It might reflect either that firms choose large-scale projects to apply for funding or that the subsidy increases the project size.

To summarize, we find that patent applications are attached to a higher likelihood of public funding if they (i) have a \textit{higher expected private value} as measured by triadic patent applications, (ii) are \textit{more general}, i.e. they cause follow-up inventions in a wide variety of different technological areas, but at the same time \textit{less original} as they more often cite backwards (iii) are \textit{co-applied}, thereby signalling more substantial knowledge exchanges through R&D collaboration, and (iv) have a \textit{larger project size} as indicated by larger inventor teams. The differences found in sample 1 thus seem to reflect the orientation of thematic R&D programs to support collaborative R&D projects (more co-applications) in technological areas of future importance (higher private value) with relevant research goals for society (more general) and a likely economic utilization of the innovation (less basic/original). Our results for sample 2 reveal that except for generality these differences remain significant once we restrict our control group to patent applications from the same technological area and time period filed by the same firm as the subsidized patent application(s) and control for firm-specific effects.

Our results point towards significant differences in the quality of patent applications from subsidized and non-subsidized R&D projects within subsidized firms. This contradicts findings of output additionality studies at the firm level for Germany which provide evidence that privately financed and publicly induced R&D show a positive and similar productivity in terms of the probability to file a patent and share of sales with market novelties (Aschhoff 2009, Czarnitzki and Hussinger 2004, Czarnitzki and Licht 2006). But it is consistent with Czarnitzki and Lopes-Bento (2014) who find that patents from subsidized firms receive on average more citations compared to patents of non-subsidized firms. Our findings elucidate that patent quality differs in various dimensions between subsidized and non-subsidized R&D projects within subsidized firms.

6.3 \textit{Robustness of results}

In this section we discuss the results of several checks to test the robustness of our results.

6.3.1 \textit{Low and high patenting firms}

We distinguish between high and low patenting firms, because firms in our sample differ strongly regarding their patenting activity. We define a firm to be low patenting if it files not more patents than
the median firm in the respective sample.\(^8\) For low patenting firms we find no significant differences between patent applications from subsidized and non-subsidized R&D projects. Certainly, this is due to the fact that the patent portfolio of low patenting firms is less diverse than the patent portfolio of high patenting firms. For high patenting firms we find that triadic applications, co-applications, citing prior art, and a larger inventor team size are correlated to a higher likelihood of being subsidized. Hence, the differences we find between patent applications from subsidized and non-subsidized R&D projects for the full sample 1 are driven by firms with a larger and more diverse patent portfolio.

6.3.2 Validity of control group

We may not identify all patent applications that stem from subsidized R&D projects in our sample for four reasons. First, we should theoretically observe all patent applications from projects that received federal funding since the collateral clauses force subsidy recipients to indicate that a patent stems from a subsidized project when seeking patent protection at the GPO. However, not all firms may have complied with that rule. It is likely though that firms which complied once are more likely to comply for other patent applications that result from subsidized projects as well. Hence, we are confident that this plays a minor role in our sample, since we only include firms that filed at least one application declaring it to be from a subsidized project.

Second, patents labelled as non-subsidized may stem from projects that received other than thematic R&D funding by the federal government. Other funding sources may be regional R&D support schemes by one of the 16 German state governments, the European Commission (EC), as well as generic R&D programs of the federal government. Regional support by German state governments plays only a minor role in public R&D funding and is more targeted at investments into R&D infrastructure.\(^9\) Similarly, generic R&D programs target mainly SMEs and usually involve no high-tech projects but enhance a firm’s R&D infrastructure and promote networks for enhanced knowledge exchange. Hence, it is unlikely that our control group is contaminated by patent applications from regional or generic R&D funding. However, there are certainly patent applications classified as non-subsidized which resulted from projects receiving support by the EU.

Third, the GPO may have missed to indicate in their online database that a patent resulted from a subsidized project. At the point the GPO added this piece of information to its online database in 2011/2012, the physical patent file including the special form on the “Information on the Funding Project Number for Intellectual Property Applications” could have already been deleted. Personal talks with the GPO reveal that in general physical patent files, though still in the online database, are destroyed approximately five years after the decision that a patent has not been granted or the applicant did not pay renewal fees. The decision to grant a patent takes on average 2 to 2.5 years at GPO.\(^10\) Hence, the information on subsidized patents is presumably more accurate in recent years in our

\(^8\) The median firm in sample 1 filed 34 applications in 1995-2005, while the median firm in sample 2 filed 16 applications.

\(^9\) Regional support also distributes R&D funding of the EU structural funds.

\(^10\) This is true if the applicant requests examination within 4 months after the application. In extreme cases the granting process may take much longer as applicants have up to 7 years to file their request of examination (http://www.dpma.de/patent/faqs/index.html#a9, accessed on June 16, 2015)
Comparing the share of patents from subsidized R&D projects, we indeed find the proportion to be higher in the recent period 2000-2005 (0.8%) than in the overall period 1995-2005 (0.7%). The difference is rather marginal though. We perform a robustness check to see whether our results are stronger for a subsample of patent applications filed between 2000 and 2005. We find the results for sample 1 to be robust in the smaller sample with the exception of COAPP which becomes insignificant. As expected, marginal effects increased in size for TRIADIC, GENERAL, NOBACK, and INVCNT.

Fourth, one may worry that patents filed after the completion of the subsidized project are not indicated as such. This is, however, not the case. For 770 subsidized patent applications from subsidized R&D projects we could link the patent application to the project-level subsidy information contained in the PROFI database. When analysing the time of patent application compared to the duration of the project, we find that 30% of the subsidized patents are filed during the project time, an additional 28% in the year of project completion, and 42% after the project ended. Among those patents filed after the project ended, 56% were filed one year after the project ended, 34% two years after the project ended, and 10% three or more years after the project completion. We therefore conclude that our database well reflects patents filed after the subsidized project ended.

Given the four potential sources, we estimate the magnitude of the sampling issue in our database to assess how it may impact our results. We use information from the PatVal survey for Germany (PatVal-EU 2005) which contains information on a sample of 3,346 German patents with a priority year between 1993 and 1997. Approximately 1.9% of those patents stem from publicly funded R&D projects. Our database differs from the data in PatVal because we only observe patent applications from firms that filed at least one patent from a subsidized project. Assuming that PatVal figures are representative, we estimate that the actual share of subsidized patents in our database should be between 2.7-3.3%. In contrast, in sample 1 we observe only 0.7% of the patent applications to result from subsidized projects. This implies that approximately 2.0-2.6% of patent applications in our sample are wrongly assigned to a non-subsidized R&D project, while they stem from a subsidized R&D project. Given that the share of public expenditure in industrial R&D in Germany dropped continuously from 10% in 1995 to 4% in 2005, it is likely that the share of subsidized patents decreased accordingly. Consequently, the share of wrongly assigned patents of 2.0-2.6% almost certainly overestimates the true share in our sample period 1995-2005. The share of patents in our control group that are potentially wrongly assigned to be a control and no subsidized patent application, is with less than 2.0-2.6% relatively small when compared to the total size of the control group which comprises 99.3% of the observations. We therefore argue that the sampling issue should not seriously impact our results.

6.3.3 Spillovers from subsidized patents to control group patents

Patent applications in the control group may be affected by subsidized patents through spillovers. In order to examine the impact such spillovers might have on our results, we excluded patents from the control group that cited a subsidized patent. We find that our results are also robust in this regard.

11 The PROFI database is a database from the Federal Ministry of Education and Research (BMBF), which contains all subsidized R&D projects from thematic R&D programs from 1969 on.
Estimation results of sample 1 with and without controlling for firm-specific effects are almost identical to the results displayed in Table 4.

In addition, we do a special robustness test for sample 2. Recall that control patents in sample 2 are from non-subsidized R&D projects filed in the same technological area and a five year window around the filing date of the subsidized patent application(s) filed by the same firm. Control patents filed in t+1 or t+2 and in the same technological area from the same firm might be affected by spillovers from the focal subsidized patent. We therefore checked how results change if we use a sample of control patents that are filed only in year t, t-1, or t-2 and the same technological area as the focal patent. Again we find that results for this new sample 2 are robust. We lose, however, almost 9,000 observations and 10 firms compared to the original estimates of sample 2 presented in section 6.2.

7. Conclusion

Little is known about differences between subsidized and non-subsidized R&D projects. In this paper we make use of a unique dataset which comprises the outcome of subsidized and non-subsidized R&D projects in terms of patent applications of firms that received a subsidy in Germany in the period 1995 to 2005. R&D projects that received public funding may differ from those that were purely privately financed for three reasons. First, firms select R&D projects according to the application criteria of the funding agency. Second, federal government selects which R&D projects it subsidizes according to their guidelines. Third, the subsidy itself may affect the quality of the project. We contribute to the literature by exploring if successful subsidized and non-subsidized R&D projects differ regarding four characteristics of the resulting patented inventions: private value, knowledge spillovers, basicness, and project size. The special design of our dataset, which comprises only firms that filed at least one application from a subsidized R&D project, allows us to control for unobserved firm heterogeneity. As all firms in our sample are subsidy receivers, we do not have to worry about selection bias arising from a firm’s selection into receiving a subsidy.

We find that subsidized R&D projects have a higher private value, are more general, less original, more often co-applied, and have a larger inventor team when compared to all other non-subsidized R&D projects by the same firms. Except for generality, these differences between subsidized and non-subsidized projects persist when limiting the control group to patent applications filed in the same technological area and application period as the subsidized patent applications of the same firm. Thus, even within the same firm and technological field, subsidized and non-subsidized projects that result in patent applications do differ. These results should be interpreted as correlations since we cannot rule out a causal subsidy effect on patented outcomes. Our findings are in contrast with results of output additionality studies conducted in a similar time period in Germany which demonstrate that publicly induced R&D expenditure shows a positive but similar productivity as privately induced R&D in terms of the likelihood to file a patent and sales share with market novelties.

12 In addition to the quality-enhancing effect, it may be that the subsidized patent would not exist without the subsidy.
In our view the differences between subsidized applications compared to all other patent applications filed by the same firm reflect the funding criteria of thematic R&D programs. German federal R&D programs are focused on the support of R&D collaborations and projects with research goals relevant for society and a likely economic utilization of results in specific technological areas which are expected to increase the German competitiveness in the future.

Surprisingly, we do not observe higher external knowledge spillovers for the outcomes of subsidized R&D projects. We leave it to future research to investigate whether this is due to the noisy measurement of external knowledge spillovers using external citations, funding agency’s difficulties to identify a project’s knowledge spillovers ex ante, disincentives in the selection of projects with high knowledge spillovers by the firm and the agency, or because spillovers are realized through unobserved channels like an increase in consumer surplus. In addition, the fact that patent applications originating from subsidized R&D projects cite more often backwards, i.e. they are less original, may reflect that the federal government supports less fundamental research projects that more likely have an economic utilization. The question is, if such required economic utilization is socially desirable.

Finally, our results hold for a variety of robustness checks including a different estimation method, sample splits, different control groups and a shorter time period. Note that the results significantly change once we include firm fixed effects which highlights the importance of controlling for unobserved firm heterogeneity in a project-level study.

A limitation of our study is that we only observe R&D projects that resulted in a patent application. It may well be that subsidized R&D projects result less often in a patent application than other projects of the firm, if the subsidized projects have a higher technological and economic risk. We leave it to future research whether our results also apply if characteristics of R&D projects are measured with other than patent statistics. In addition, strictly speaking we cannot provide causal evidence due to the data characteristics. The question which differences between subsidized and non-subsidized R&D projects are due to the selection efforts (of the firms and of the funding agency) or the effect of the subsidy on the project outcome calls for further research.

References


