ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Koehler, Mila; Peters, Bettina

Working Paper

Subsidized and non-subsidized R&D projects: Do they differ?

ZEW Discussion Papers, No. 17-042

Provided in Cooperation with:
ZEW - Leibniz Centre for European Economic Research

Suggested Citation: Koehler, Mila; Peters, Bettina (2017) : Subsidized and non-subsidized R&D projects: Do they differ?, ZEW Discussion Papers, No. 17-042, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim

This Version is available at:
https://hdl.handle.net/10419/170583

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Discussion Paper No. 17-042

Subsidized and Non-subsidized R&D Projects: Do They Differ?

Mila Koehler and Bettina Peters

ZEW
Zentrum für Europäische Wirtschaftsforschung GmbH
Centre for European Economic Research

Discussion Paper No. 17-042

Subsidized and Non-subsidized R&D Projects: Do They Differ?

Mila Koehler and Bettina Peters

Download this ZEW Discussion Paper from our ftp server:

http://ftp.zew.de/pub/zew-docs/dp/dp17042.pdf

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the ZEW.

Subsidized and non-subsidized R&D projects: Do they differ?*

Mila Koehler[a, b] and Bettina Peters[a,c,d]

a) Centre for European Economic Research (ZEW), Department of Innovation Economics and Industrial Dynamics, Mannheim, Germany
b) KU Leuven, Department of Managerial Economics, Strategy and Innovation, Leuven, Belgium
c) Mannheim Centre for Competition and Innovation (MaCCI), Mannheim, Germany
d) University of Luxembourg, Luxembourg

October 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 within the same firm 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 collaborative R&D projects that have an immediate economic utilization of results.

JEL: H25, H50, O31, O34, O38
Keywords: R&D, subsidies, patents, evaluation, project-level analysis, within firm comparison

Contact information

Mila Koehler
Centre for European Economic Research (ZEW), Mannheim, Germany
Postal Address: Centre for European Economic Research (ZEW), L7 1, 68161 Mannheim, Germany
Phone: +49 621 1235 283, e-mail: mila.koehler@zew.de
Bettina Peters
Centre for European Economic Research (ZEW), Mannheim, Germany
Postal Address: Centre for European Economic Research (ZEW), L7 1, 68161 Mannheim, Germany
Phone: +49 621 1235 174, e-mail: b.peters@zew.de

---

* Acknowledgements: We would like to thank the German Patent and Trademark Office and Thorsten Doherr for their help in creating the database. Furthermore, we are grateful to Dirk Czarnitzki, Pierre Mohnen, Reinhilde Veugelers, Otto Toivanen, Georg Licht, Christian Koehler, Cindy Lopes-Bento, Michal Pellens, Vivek Bhattacharya, Markus Simeth and Aloña Martiarena for helpful comments and discussions. We also thank participants at the ZEW Innopat Conference (Mannheim 2015), EPIP (Oxford 2016), IIOC (Boston 2017) and DRUID (New York 2017) for helpful comments. We gratefully acknowledge financial support from the international research program "Strengthening Efficiency and Competitiveness in the European Knowledge Economies" (SEEK) and the Flemish Science Foundation (FWO aspirant scholarship). Any errors remain our own.

1. Introduction

Most governments worldwide aim at spurring private R&D investments by using R&D subsidies or R&D tax credits. The main argument for both policy measures is a suboptimal level of R&D due to market failure rooted in spillovers and financial constraints. In the EU for example the share of government spending on total annual R&D expenditure ranged from 12% to 14% in the years 2000 till 2015 (Eurostat 2017). Such massive public support raises at least three important questions: First, do public R&D subsidies crowd out private R&D investments? Second, is publicly financed R&D as productive as privately financed R&D, e.g. in terms of number of patents? Third, which types of R&D projects receive public funding and do subsidized R&D projects differ from purely privately financed ones? The literature has mainly focussed on the first two questions by evaluating 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 about the same (see e.g. Czarnitzki and Hussinger 2017). All these studies, however, use firm-level data and thus fail to answer the third question: which types of R&D projects receive funding within firms, and whether and in which dimensions these projects differ from non-subsidized ones.

That not much is known about differences between R&D projects that are subsidized and those that are not, 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. 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). Public support by the federal government to industrial R&D ranged between 1-1.3 billion euro per year in the period of consideration 1995-2005, which corresponded to 4-10% of total annual industrial R&D expenditures (BMBF 2010).

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. Surprisingly, subsidized R&D projects are not associated with larger external knowledge spillovers. We argue, that our findings reflect the orientation of German federal thematic R&D programs[1] to support collaborative R&D projects with a likely economic utilization of the innovation aiming to develop basic technologies in promising high-tech fields.

[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)".

Most interestingly, except for generality these differences remain even 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. Our findings are enlightening on results of output additionality studies conducted in a similar time period in Germany. They have demonstrated that publicly induced R&D expenditure shows a positive and 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 2017, Czarnitzki and Licht 2006). The latter studies, however, are conducted on the firm level, i.e. they cannot distinguish which patents a firm filed from a subsidized and which from a non-subsidized R&D project, highlighting the importance to study effects of R&D subsidies on the project, not the firm level.

This paper is structured as follows. Section 2 briefly explains why governments distribute R&D subsidies and 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. Section 6 presents descriptive and estimation results. Section 7 concludes.

2. Public support for R&D in Germany

Given the crucial role of R&D for competitiveness and growth (see e.g. the survey by Hall et al 2010) most other governments worldwide support 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 of 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.

In Germany, R&D subsidies to the industry are distributed either via thematic or generic R&D programs (BMBF 2012). Thematic R&D programs (“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 (“Indirekte Projektförderung”) are not

directed to specific technological areas. These programs are especially targeted on SMEs and firms located in East 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 management 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 help of the scientific experts at the management 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 1986, EC 1996, EC 2006). It foresees that the subsidy amount may at most cover 50% of the total project costs. A mark-up may be paid for SMEs, collaborative projects and firms located in East 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 selected technological fields.

3. Literature review and contribution

3.1 Literature review

There are only a few existing studies dedicated to analysing differences between subsidized and purely privately financed R&D projects. This is mainly due to the fact that project-level data are hardly available. To the best of our knowledge the only studies in this respect 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 PatVal[2] 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 as patents from subsidized R&D projects.

Takalo et al. (2013a) analyse the funding decision of the granting agency in Finland. For a sample of 379 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 and lower in terms of their

[2] The PatVal database comprises information on a sample of European inventors holding a patent at the European Patent Office (EPO).

expected economic risk than projects of rejected applicants. The latter effect is only weakly significant. The authors make use of project-level information of submitted R&D projects, while they do not observe any R&D projects firms did not seek public support for. Though interesting, the findings of Schneider (2008) and Takalo et al. (2013a) still leave it open whether firms choose to apply with more innovative R&D projects. So far, nothing is known about how subsidized R&D projects differ in their outcomes from non-subsidized R&D projects conducted by the same firm.

Our work is also related to the vast evaluation literature which analyses the effect of R&D subsidies at the *firm-level*. One major concern is that public R&D subsidies crowd out private R&D investments, i.e. private R&D expenditures are 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 these evaluation studies are confronted with is accounting for the selection bias in order to identify the causal effect of the subsidy. A selection bias occurs if especially more 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. Studies analysing differences between subsidized and non-subsidized firms indeed find that subsidized firms are more successful. In Germany, for instance, firms are more likely to receive R&D funding if they, conduct R&D continuously, have more innovative capacities, already received public funding in the past, have better trained personnel, and compete on international markets, but are not foreign owned (see e.g. Aschhoff 2010, Hussinger 2008, and Czarnitzki and Lopes-Bento 2014).

Most input additionality studies focus on testing the full crowding out hypothesis, i.e. they investigate if subsidized firms spend more on R&D than in the counterfactual situation. The majority of studies concludes that subsidized firms spend more on R&D than in the counterfactual situation when not receiving a subsidy (for an overview see the surveys of Zúñiga-Vicente et al. 2014, Klette et al. 2000 or David et al. 2000; and for studies on Germany see Czarnitzki and Fier 2002, Almus and Czarnitzki 2003, Czarnitzki and Licht 2006, Aerts and Schmidt 2006, Czarnitzki et al. 2007, Hussinger 2008, Aschhoff 2009, Czarnitzki and Lopes-Bento 2014, Hud and Hussinger 2015, Czarnitzki and Hussinger 2017).[3] Only a few studies tested if additionality exists, i.e. if subsidized firms spend more on R&D *after* subtracting the amount of the received subsidy. Czarnitzki and Hussinger (2017), Hud and Hussinger (2015) and Aschhoff (2009) find on average additionality of public R&D subsidies. However, one should keep in mind that input additionality is not a necessary condition for a welfare-improving R&D subsidy.[4]

It is unclear though whether R&D subsidies improve 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,

[3] Only one study for Germany (Czarnitzki et al. 2007) finds full crowding out of private investments for a small sub-group 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 non-subsidized firms.

[4] As Takalo et al. (2013b) point out additionality should always be found beneficial at the extensive margin, i.e. for those supported firms which would not conduct any R&D projects in absence of the subsidy. At the intensive margin, however, projects that exhibit high spillovers to society are less likely to show additionality than projects with low spillovers. If a project exhibits large spillovers to society, the funding agency may want to support this project to a larger extent even if it crowds out private R&D investments because society profits more at the margin.

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 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. Empirical evidence so far is mixed. 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 2017, 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. It is, however, unclear if this finding is driven by patents from subsidized or non-subsidized R&D projects since firm-level studies fail to distinguish between outputs from both types of R&D projects. To disentangle the effect of publicly and privately induced R&D, *project-level* data is required.

3.2 Contribution

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 resulted in 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 with 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 is expected to 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 (see section 2). It is unclear though if agencies indeed pick projects with a high value for society through knowledge spillovers or tend to fund primarily privately profitable projects. The project selection behaviour of the agency in turn affects a firm's choice of projects to apply for subsidies. Third, there may be an effect of the subsidy on the project outcome. The subsidy may for instance enable a firm to employ better researchers, instruments or equipment or grant access to better networks and information resulting in a higher quality of the research project.

In this paper, we investigate whether subsidized and non-subsidized R&D projects that resulted in patent applications differ in several key characteristics such as their private value, knowledge spillovers, basicness and project size. We have created a novel dataset by identifying all patent applications filed in

[5] But 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 hiring and training researchers. Arqué-Castells and Mohnen (2015) 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.

Germany between 1995 and 2005 that stem from a subsidized R&D project and by merging all patent applications from non-subsidized R&D projects of the same applicant firms. 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 for unobserved firm heterogeneity because we compare patent applications from subsidized and non-subsidized R&D projects of the *same* firms. Equally important, 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. Our approach can be seen as a reduced form model of supply and demand. As we cannot rule out reverse causality, we do not intent 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

$$(1) \qquad P(SUBS_{ij}) = \Phi(X'_{ij}\beta + Y'_{ij}\beta + Z'_{j}\beta),$$

where $SUBS_{ij}$ takes value one if patent i of firm j stems from a subsidized R&D project and zero otherwise. X'_{ij} is a vector of patent characteristics that captures the expected private value, knowledge spillovers, basicness, and the size of a R&D project. Y'_{ij} is a vector of year and technological area dummies, and Z'_{j} 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 (an indicator for high private value), i.e. patents filed in the EU, Japan and US. At the same time they have 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 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 uncover the program selection. 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 program selection effect. Firms select projects from technological areas that comply with the orientation of thematic R&D programs.

5. Data and variables

Our empirical analysis is based on patent application data. The use of patent data in order to learn about firms' inventions is controversially discussed in the literature (see e.g. Cohen and Levin 1989). The major virtues of patent data are its public availability and the provision of an objective measure of an inventive step. However, not all inventions are patented either because the knowledge cannot be protected by a patent or because firms decide not to patent the knowledge for strategic reasons.[6] In addition, applying for patent protection costs time and money. This is why larger firms are more likely to protect an invention with a patent than smaller firms. Hence, our data is more representative for technological sectors where firms generally *can* patent their inventions, and for firms that *do* patent their inventions.

5.1 Construction of the patent data base

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 (NKFT88, NKBF98)[7] 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 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

[6] They may find it more profitable to keep inventions secret and use e.g. lead-time and product complexity to appropriate the returns of R&D investments (Mansfield 1985).

[7] 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.

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) 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, Table 5 in Appendix 1). As application period we define a three-year window, i.e. up to two years before 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, why we dropped these observations from sample 2. In total, sample 2 contains 61,065 patent applications filed by 206 firms of which 1,127 stem from subsidized R&D projects (1.8%).

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 14 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. For the median firm the share amounts to 4.77% which corresponds to around 20 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 and high patenting firms, as well as draw a random sample restricting the number of control patents per subsidized patent (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

	Number of patent applications per firm from…		**Share of patents from subsidized R&D projects (in %)**
	...subsidized R&D projects	...non-subsidized R&D projects	
Minimum	1	1	0.01
1st percentile	1	1	0.05
5th percentile	1	2	0.11
10th percentile	1	3	0.25
25th percentile	1	9	1.09
50th percentile	1	32	4.77
75th percentile	3	248	16.67
90th percentile	7	1,229	33.33
95th percentile	14	2,275	50.00
99th percentile	60	11,711	63.64
Maximum	287	26,458	88.89
No. of firms	249	249	249

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 investigate 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.

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 receives by subsequent patent applications filed by the same firm. A citation by patent B to patent A indicates that

patent B draws from the previously existing knowledge contained in patent A.[8] Hence, self-citations reflect follow-on 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 (Trajtenberg et al. 1997). 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.[9] Moreover, when collecting the citation data, we do not observe future citations to the patent. Consequently, this would lead older patents to be cited more often. Due to the truncation problem, 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). Hall et al. (2005) likewise stressed that a patent's prime citation years are usually the ones early in its life cycle, more precisely those in a three- to ten-year window.

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 and 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. As well-known 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. On the other hand there may be citations without spillovers. At EPO and GPO citations are included by the patent examiner who may have added content-related references that did not serve as a source of inspiration to the inventing firm. Despite these obvious flaws, forward citations have the advantage of leaving a measurable "paper trail" of an otherwise unobserved phenomenon (see Jaffe et al. 1993 for a detailed discussion of the use of forward citation to measure knowledge spillovers). In our research context, the key question for drawing valid inferences about knowledge spillovers from citations is whether we expect that (i) citations without spillovers and (ii) spillovers without citations systematically

[8] Citations to prior work are included by the patent examiner during the process of granting a patent at the GPO and the EPO. The examiner tries to disprove the novelty of a patent application by searching the web of existing patents, patent applications and non-patent literature for inventions on which the current application builds upon. Cumulative inventions are only reflected in a citation if the examiner adds the reference to the patent. The references are related to the application content-wise, but they may not have served as a source of inspiration to the inventing firm. The latter should be less of a worry for self-citations though, as previous inventions of the same firm are likely to serve as a source of inspiration for subsequent inventions within the firm.

[9] 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.

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 both groups of patent applications.

Furthermore, knowledge is intendedly exchanged in collaborative projects. *Co-application* (COAPP) is a dummy variable that equals one if there are two or more applicants. A co-application unambiguously indicates a collaboration with another firm, university or research institute. Not all collaborations, however, result in a co-application. 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.

Basicness. An invention is regarded as basic if it focuses on a scientific rather than a technological question if it has a large impact within a given field or a widespread impact across a broad range of fields (Jaffe and Trajtenberg 2002). We approximate the basicness of an invention by three indicators: the number of backward citations, originality, and generality (Trajtenberg et al. 1997).

Backward citations (BACKCIT) reflect to which extent the invention draws from pre-existing knowledge. 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 (Trajtenberg et al. 1997).

An invention is assumed to be more original or complex if it draws from many different technological areas. 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:

$$(3) \qquad Originality_i = 1 - \sum_{k=1}^{N_i} \left(\frac{Cited_{ik}}{Cited_i}\right)^2,$$

where $Cited_{ik}$ are the number of backward citations to prior art by the patent i in the 3-digit technology class k and N_i is the number of different 3-digit technology classes to which the cited patents belong. The originality index ranges between zero and one with higher values representing a wider variety of cited technological fields. Furthermore we 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.

A technology is assumed to be more basic, the more inventions it triggers in a broader range of technological fields. This idea is captured by the g*enerality* (GENERAL) index which is defined by one minus the Hirschman-Herfindahl index of the citing 3-digit technology classes:

$$(2) \qquad Generality_i = 1 - \sum_{k=1}^{N_i} \left(\frac{Citing_{ik}}{Citing_i}\right)^2,$$

where $Citing_{ik}$ is the number of forward citations within 5 years to patent i from 3-digit technology class k. 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.

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.

As the employed patent indicators are partly based on the same sources, we conducted a correlation analysis of the indicators, which is displayed in Appendix 3. We find that correlation of the indicators does not give concern to multicollinearity issues in our analysis.

6. Empirical Evidence

This section demonstrates empirical evidence on the question whether and to what extent subsidized and non-subsidized R&D projects differ. We first present descriptive statistics, followed by the findings for our baseline econometric model and several robustness tests.

6.1 Descriptive evidence

Table 2 displays summary statistics of our patent indicators in total for all patent applications and separately for subsidized and non-subsidized patent applications. The table also contains t-tests on mean differences between both groups. All summary statistics are for sample 1, but results for sample 2 are very similar (see Table 4 in Appendix 1).

On average, the patents in our sample receives 0.3 self-citations and 0.7 external citations. 30.7% of them are triadic patents. The mean number of inventors is about 2.7 and 3.2% of the patents were co-applied. The unconditional mean difference t-tests in Table 2 furthermore show that patents from subsidized and non-subsidized projects differ significantly in various but not all dimensions. Subsidized patents are less basic in the sense of having a higher likelihood of citing any prior art at all. 39.5% of subsidized patents have no backward citations implying that 60.5% cite prior art compared to 45.9% among non-subsidized patents. In addition subsidized patents cite more prior art, which is indicated by the higher number of backward citations (2.38 vs 1.86). On the other hand, subsidized patents score higher on the originality index (0.149) than non-subsidized patents (0.112) which shows that the subsidized inventions draw on prior knowledge from a broader set of technological fields. However, this comparison does not take into account that subsidized patents cite prior art more often. Therefore, we additionally run a t-test on differences in the originality index only for those patent applications that have at least one backward citation. For this subsample, we do not detect any significant differences in the originality index of subsidized and non-subsidized patents. There are likewise no significant differences regarding the generality and the likelihood to be cited at all (NOFORW) between subsidized and non-subsidized patents.

Table 2: Descriptive statistics of patent applications from subsidized and non-subsidized R&D projects 1995-2005

Variable	All				Subsidized				Non-Subsidized				t-test
	mean	s.d.	min	max	mean	s.d.	min	max	mean	s.d.	min	max	
Private value													
SELFCIT	0.306	0.994	0	47	0.358	0.972	0	9	0.306	0.995	0	47	
TRIADIC	0.307	0.461	0	1	0.294	0.456	0	1	0.307	0.461	0	1	
Spillovers													
EXTCIT	0.728	1.583	0	87	0.593	1.198	0	13	**0.729**	1.585	0	87	**
COAPP	0.032	0.177	0	1	**0.085**	0.278	0	1	0.032	0.176	0	1	***
Basicness													
GENERAL	0.048	0.140	0	0.889	0.053	0.150	0	0.778	0.048	0.140	0	0.889	
NOFORW	0.560	0.496	0	1	0.572	0.495	0	1	0.560	0.496	0	1	
ORIGINAL	0.112	0.210	0	0.960	**0.149**	0.239	0	0.878	0.112	0.210	0	0.960	***
BACKCIT	1.862	2.751	0	44	**2.381**	2.972	0	28	1.858	2.749	0	44	***
NOBACK	0.540	0.498	0	1	0.395	0.489	0	1	**0.541**	0.498	0	1	***
Project size													
INVCNT	2.728	1.898	1	26	**3.017**	1.837	1	14	2.726	1.898	1	26	***
No. of obs.	174,311				1,207				173,104				
No. of firms	249				249				249				

Notes: All values are calculated using sample 1. *p<0.10, **p<0.05, ***p<0.01. If the difference between patents from subsidized and non-subsidized R&D projects is significant above the 5%-level, we bold printed the higher of the two means.

The evidence for *knowledge spillovers* is also mixed. We find a significantly higher share of co-applications among subsidized patents (8.5%) than among non-subsidized patents (3.2%) indicating a higher degree of knowledge exchange. This difference 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.[10] Most interestingly and in contrast to Schneider (2008), our within-firm comparison reveals significantly lower external citations for subsidized 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. However, this result might also be driven by underlying differences in technology, time or other firm characteristics which calls for a multivariate analysis. Indeed this view is supported by sample 2 for which we do not find any significant differences between both groups (see Table 3).[11]

Furthermore, the descriptive evidence in Table 2 demonstrates no significant differences with respect to *private value* indicators, i.e. self-citations and triadic patent applications. The *inventor team* is significantly

[10] For details see Veugelers (1998) who reviews theoretical and empirical findings on R&D collaborations.

[11] Recall that in sample 2 our control patents stem from the same technological area and application period as the subsidized patents filed by the same firm.

larger for patent applications from subsidized R&D projects. It is unclear though if this finding reflects that firms apply with larger projects or if the subsidy increases the project size.

6.2 Multivariate analysis

We estimate the Probit model on SUBS for sample 1 and 2 (see equation 1). We are well aware of potential endogeneity concerns due to reverse causality. 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 instance enables firms to employ higher qualified researchers or better technical equipment.[12] Since the set of information for non-subsidized projects is limited to patent characteristics, we could not find suitable instruments to follow an IV approach. We therefore interpret all following results as correlations and not in a causal way. Still, the multivariate Probit analysis allows us to learn in which characteristics subsidized and non-subsidized projects differ controlling for firm, technology and year effects.

Table 3 displays the average marginal effects of the Probit estimations for sample 1 (Columns 1 and 2) and sample 2 (Columns 3 and 4) each with and without controlling for firm heterogeneity. For ease of presentation, displayed marginal effects and standard errors are multiplied by 100. 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 R^2 and the count R^2. All in all, the model seems to fit the data quite well. Based on the preferred specifications including the firm dummies, the McFadden's pseudo R^2 varies between 23.0 and 19.7 and the model correctly predicts the outcome for 77.1% and 75.6% of the observations in both samples.

Based on sample 1, our results show that 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 (Column 2). Recall that a triadic patent application indicates a higher private 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 (Column 1) which highlights the importance to account for firm-specific effects.

Looking at the average marginal effects of the Probit at Table 3, it strikes that they are very low. But keep in mind that the share of subsidized R&D projects is very low with 0.7% in sample 1 and 1.8% in sample 2. Let us consider for example the effect of a triadic patent application. If a patent is a triadic patent application, this is associated with an increase in the proability that the patent stems from a subsidized project by 0.13 percentage points in sample 1. Considering that the share of subsidized patents in sample 1 is 0.7% only, this corresponds to a 19% increase.

While triadic patents can be seen as measuring the expected direct benefits of the inventions, self-citations measure indirect benefits due to own follow-on inventions. The results do not show, however, a significant correlation between the number of self-citations (SELFCIT) and the probability of receiving a subsidy. Our results suggest that firms choose projects to apply for public funding based on the expected

[12] In addition to the quality-enhancing effect, it may be that the subsidized patent would not exist without the subsidy.

direct benefits of the invention without taking future indirect benefits, as measured by self-citations, into account.

Table 3: Average marginal effects of the Probit estimation on SUBS, 1995-2005

Variable	Sample 1		Sample 2	
	(1)	(2)	(3)	(4)
Private value				
SELFCIT	0.029	0.016	0.048	0.015
	(0.021)	(0.020)	(0.059)	(0.057)
TRIADIC	-0.034	0.128**	0.274**	0.462***
	(0.046)	(0.051)	(0.134)	(0.142)
Spillovers				
EXTCIT	-0.038*	-0.017	-0.004	-0.024
	(0.020)	(0.018)	(0.051)	(0.051)
COAPP	0.869***	0.413***	2.382***	0.895**
	(0.158)	(0.129)	(0.496)	(0.369)
Basicness				
GENERAL	0.326**	0.345**	0.498	0.712
	(0.166)	(0.159)	(0.461)	(0.447)
NOFORW	-0.030	-0.066	-0.080	-0.092
	(0.051)	(0.049)	(0.136)	(0.131)
ORIGINAL	-0.005	-0.006	-0.030	0.078
	(0.110)	(0.107)	(0.136)	(0.291)
BACKCIT	-0.008	0.002	0.014	-0.035
	(0.010)	(0.010)	(0.030)	(0.029)
NOBACK	-0.448***	-0.271***	-0.871***	-0.770***
	(0.061)	(0.058)	(0.162)	(0.159)
Project size				
INVCNT	0.044***	0.075***	0.151***	0.188***
	(0.011)	(0.011)	(0.030)	(0.030)
Year dummies	Yes	Yes	Yes	Yes
Area dummies	Yes	Yes	Yes	Yes
Firm dummies	No	Yes	No	Yes
McFadden's R^2	0.079	0.230	0.075	0.197
McKelvey & Zavoina's R^2	0.145	0.261	0.117	0.221
Count R^2	64.97%	77.11%	65.58%	75.64%
No. of obs.	174,311	174,311	61,065	61,065
No. of firms	249	249	206	206

Notes: Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01. Marginal effects, coefficients and standard errors are multiplied by 100 for ease of representation. 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.

Our finding that projects with a higher expected private value are associated with a higher probability of receiving a subsidy might be driven by a program selection effect. Thematic R&D programs in Germany are targeted at specific technological areas which are more likely to provide higher profit opportunities and thus higher private values of resulting inventions. In order to control for a thematic program selection effect, we use sample 2 which compares subsidized and non-subsidized patent applications filed by the same firm within the same technological area and application period (Column 3 and 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 or because the subsidy itself increases the quality and hence the value of the project.

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-on inventions by other firms. However, the econometric evidence for our two spillover measures 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 effort of the agency to support knowledge exchange by fostering R&D collaborations results into true collaboration.

Surprisingly though patent applications with higher expected knowledge spillovers – as measured by external forward citations (EXTCIT) – are not correlated with a higher likelihood of getting public funding. 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. This result remains valid in sample 2 when we compare technologically similar patents. The question remains open for future research if our finding is due to disincentives to firms to apply with more socially valuable projects, disincentives to the agency to select projects with especially high knowledge spillovers, difficulties to judge a project's knowledge spillovers to society ex ante, and/or the noisy measure of external knowledge spillovers. By noisy measurement, we mean on the one hand that examiner-added external citations might not reflect spillovers and one the hand that spillovers might be realized through other unobserved channels like an increase in consumer surplus. One reason for a disincentive to choose projects with high knowledge spillovers on the firm and the agency side may be that the German funding agency promotes projects which have a likely economic utilization. Projects with a likely economic utilization are less fundamental, as they are close to be marketed and hence associated to a lower risk of invention.

Our finding is, however, in contrast with Czarnitzki and Lopes-Bento (2014) who find that patents from subsidized firms receive on average more citations compared to patents from non-subsidized firms. Czarnitzki and Lopes-Bento (2014) cannot, however, differentiate between patents from subsidized and non-subsidized projects. They only differentiate between subsidized and non-subsidized firms accounting for potential selection bias employing a matching technique. Hence, our result emphasises the importance of conducting project-level analysis.

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-on 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, our results speak against reverse causality since we should observe higher external forward citations in this case. Therefore, our findings are rather affected by firm and agency selection than by the subsidy itself. The finding suggests that the German funding agency is able to identify future areas with higher generality, i.e. of higher value for society.

Like in the descriptive statistics, we corroborate 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 patented inventions from subsidized R&D projects are more general affecting future inventions in a wider variety of technologies. This result seems to be driven by thematic preferences of agencies. On the other hand 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 whether the required economic utilization is socially desirable. From a theoretical point of view, fundamental research projects are more likely to suffer from financial constraints due to the high uncertainty of an outcome that may be economically utilized.

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. Finding that patents from subsidized projects are invented by an on average larger team may indicate that the additional money available through the subsidy enables firms to employ more researchers, thereby increasing the quality of the research project. It may also reflect that firms tend to apply for public funding with larger, more personnel intensive projects, which the firm otherwise would maybe not be able to finance itself.

While output additionality studies have demonstrated 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 2017, Czarnitzki and Licht 2006), our findings elucidate that the patent quality within subsidized firms differs in various dimensions between subsidized and non-subsidized R&D projects. To summarize, controlling for firm specific effects we find that patent applications are attached to a higher likelihood of public funding if they (i) have a *higher expected private value* as measured by triadic patent applications, (ii) are *more general* in sample 1 but not 2, i.e. they cause follow-on inventions in a wide variety of different technological areas, (iii) but are at the same time *less original* as they more often cite backwards, (iv) are *co-applied*, thereby signalling more substantial knowledge exchanges through R&D collaboration, and (v) have a *larger project size* as indicated by larger inventor teams.

In our view the findings suggest that the German funding agency is well suited to select projects that induce knowledge spillovers through collaboration in technological areas promoting more general application of the resulting inventions. The subsidy itself may even stimulate firms to conduct larger research projects, which may have not been conducted in the absence of the subsidy. In contrast, the

focus of the German funding agency on selecting those projects whose research results have a likely economic utilization deserves in our view a more critical discussion. Theoretically speaking, would it be socially desirable to finance especially those projects with high knowledge spillovers to society. Our research indicates though that the current focus on projects with a high economic utilization may lead to the support of projects that result in inventions of lower originality and higher private value, while not inducing more knowledge spillovers via external citations when compared to patented inventions from non-subsidized projects.

6.3 Robustness of results

We perform several checks to test the robustness of our findings. For simplicity we only present robustness results for the model including firm specific effects. Estimations results are shown in Table 6 (sample 1) and Table 7 (sample 2) in Appendix 1.

Heteroscedasticity-robust standard errors. As heteroskedastic standard errors lead to biased coefficients in the Probit model, we first check if our results hold when estimating a linear probability model using an OLS regression with heteroscedasticity-robust standard errors. Column 2 in Table 6 and Table 7 confirm our prior results for both samples.

Low and high patenting firms. We distinguish between low and high patenting firms because firms in our sample differ strongly regarding their patenting activity (see section 5.2). We define a firm to be high patenting if it files at least as much patents as the firm at the 90th percentile in the respective sample.[13] For low patenting firms we find for both samples co-applications, citing prior art, project size, and for sample 1 also scoring higher on the generality index to be correlated to a higher likelihood of being subsidized (see Column 3 of Table 6 and Table 7). For high patenting firms we find for both samples a positive correlation of triadic patent applications, citing prior art and project size with the likelihood to get publicly funded.[14] The positive correlation between triadic patent applications and the probability to receive a subsidy seems driven by high-patenting firms. This suggests that especially high patenting firms tend to apply for public funding with projects that have a higher private value.[15]

Regarding the number of co-applications the sample split reveals that co-applications are significantly positively correlated with the probability to receive public funding only for low patenting firms. The question is, if this result reflects that the effort of the agency to support knowledge exchange between collaborating partners is more fruitful for low patenting firms. It may also picture that low patenting firms shoulder the higher costs of co-applying in terms of managing the legal complexities between IP rights

[13] The firm in the 90th percentile filed 1231 patent applications in sample 1 and 262 applications in sample 2 in the years 1995 until 2005.

[14] We also analysed if the size of the effects differs for low and high patenting firms. For that purpose, we estimated a Probit model for sample 1 and 2 accounting for firm-specific heterogeneity including an interaction effect of our patent characteristics with a dummy for high patenting firms. The interaction effects are not significant at the 5%-level except for INVCNT x HIGH which is significantly negative in both samples. Hence, for high patenting firms projects with a smaller size that result into a patent application are correlated with a higher likelihood of being subsidized compared to projects of low patenting firms.

[15] In this case a subsidy effect seems improbable. We find it rather unlikely that the subsidy should increase the quality and hence the private value of projects of high patenting firms, while not affecting the quality of low patenting firms.

across partners more likely in case they receive a subsidy. In addition, high patenting firms may have more efficient strategies to file patents from a collaborative project alone in order to avoid costly co-managing of IP rights.[16] Besides, the finding on generality suggests that the program selection effect we find is not driven by the 10% highest patenting firms. This may be explained by the larger patent portfolios of high patenting firms.

Random draw of control patents. As the share of subsidized patents is with 1.8% of subsidized patents in sample 2 still very low, we check if our results hold when reducing the number of control patents per subsidized patent. For firms that have more than five control patents per subsidized patent, we randomly draw 5 control patents per subsidized patent in sample 2. Running 200 iterations, we find that the results confirm our prior findings (see Column 6 of Table 7).

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 (see Column 4 of Table 6 and Table 7). Estimation results of sample 1 and 2 controlling for firm-specific effects are almost identical to the results displayed in Table 3.

Validity of control group. Readers may raise the concern that we do not identify all patent applications that stem from subsidized R&D projects in our sample. Indeed, we identify four reasons for a potential misassignment of subsidized and non-subsidized patent applications in our sample. In the following we discuss each reason and how it may affect the control group patents estimation results below. Our overall conclusion from this exercise is that misassignment should not seriously affect our results.

First, in theory we should 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. Admittedly, not all firms may have complied with that rule. But firms which complied once are well aware of this rule and hence are very likely to comply for other patent applications that result from subsidized projects. This is especially the case within the same application period, i.e. for sample 2. Since we only include firms in our sample that filed at least one subsidized patent application, we are confident that this plays a minor role in our sample.

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. They are targeted mainly at investments into SMEs, R&D personnel and R&D infrastructure. This has also to do with the fact that money from the EU structural funds under the EU cohesion policy, i.e. European Regional Development Fund (ERDF) and European Social Fund (ESF), is distributed by the German state governments. Guidelines for ERDF and ESF foresee to support R&D infrastructure, employment and knowledge exchange through network activities with a strong focus on

[16] We checked if our finding is driven by the fact that high patenting firms also file more co-applications when not receiving a subsidy. This, however, is not the case. The share of co-applications for high (low) patenting firms is among non-subsidized patents 2.9% (4.9%) and among subsidized patents 3.9% (13%) in sample 1. Results are similar for sample 2.

SMEs (EC 1993a, EC 1993b, EC 1999a, EC 1999b). Similarly, generic R&D programs are usually not directed to specific technological areas, target mainly SMEs and allow to support R&D infrastructure, research cooperation, and innovation networks (BMBF 2012). Hence, it is unlikely that our control group is contaminated to a larger extent by patent applications from regional or generic R&D funding. Regarding the EU framework program most of the key actions focus on enhancing R&D infrastructure through channels like EU-wide networking and training activities, as well as fostering international knowledge exchange and international collaboration.[17] The framework program supports projects in specific technological areas like e.g. life science, environment and IT at research institutions, universities as well as enterprises. So when it comes to the support of high-tech R&D projects directly by the EU, we might indeed miss patent applications in our database. As the framework program money is, however, distributed EU-wide to international projects and many key actions of the framework program are indirect support through service networks and training, we expect only a small bias through this channel.

Third, the GPO may have missed to indicate in their online database that a patent resulted from a subsidized project. The GPO added this piece of information to its online database in 2011/2012. At this point 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.[18] Hence, the information on subsidized patents is presumably more accurate in recent years in our sample. Comparing the share of patents from subsidized R&D projects in sample 1, 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. Column 5 in Table 6 and Table 7 corroborates our previous results for sample 1 and 2 with the exception of COAPP which becomes insignificant (Column 5 of Table 6 and Table 7). As expected, marginal effects increased in size for TRIADIC, NOBACK, INVCNT and GENERAL. The latter even becomes significant in sample 2.

Fourth, one may worry that patents filed after the completion of the subsidized project are not indicated as such. For a subsample of 770 subsidized patent applications we could link the patent application to the project-level subsidy information contained in the PROFI database. The PROFI database is a database from the Federal Ministry of Education and Research (BMBF), which contains information on all subsidized R&D projects from thematic R&D programs from 1969 on. When comparing the first filing date of the patent application with the duration of the subsidized R&D 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 subsidized project ended, 56% were filed one year

[17] http://cordis.europa.eu/fp5/, http://cordis.europa.eu/fp5/src/struct.htm#KEYACTIONS and http://cordis.europa.eu/fp6/, accessed on October 9, 2017.

[18] 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)

after, 34% two years after, 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.

Though we are confident that the aforementioned four potential sources of misclassification of subsidized and non-subsidized patents should not affect our results to a large extent, we additionally tried to estimate the magnitude of the measurement error in our database and assess how it may impact our results. We use information from the PatVal survey for Germany (PatVal-EU 2005) which contains data 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% (see Appendix 2 for details on the calculation of the measurement error). 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. However, even assuming a measurement error of 2.0-2.6% in the group of control patents is relatively small when compared to the total size of the control group which comprises 99.3% of the observations. We therefore argue that the measurement error should not seriously impact our results.

7. Conclusion

So far little is known about within-firm 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 R&D 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 patent 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. Our analysis additionally controls for a thematic program selection effect. Since we cannot rule out a causal subsidy effect on patented outcomes our results should only be interpreted as correlations.

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 within-firm 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. Thus, even within the same firm and technological field, subsidized and non-subsidized projects that result in patent applications do differ. Our results are in contrast to results of output additionality studies on the firm level conducted in a similar time period in Germany which find 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. This demonstrates the importance to conduct analysis of the effect of R&D subsidies on the project not only the firm level.

In our view the differences between subsidized applications compared to all other patent applications filed by the same firm reflect partly 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. Our findings suggest that the German funding agency selects projects that induce knowledge spillovers through collaboration in technological areas promoting more general application of the resulting inventions. The subsidy itself may even induce larger research projects, which may have not been conducted in the absence of the subsidy. In addition, we find that subsidized patents cite more often backwards, i.e. are less original, and that they are not associated with higher external knowledge spillovers. This may reflect that the federal government supports less fundamental research projects which are more likely to have an economic utilization. The question remains if such required economic utilization is socially desirable. From a theoretical point of view, it would be socially desirable to finance especially those projects with high knowledge spillovers to society. We furthermore rule out that the higher private value of subsidized patents is due to technology preferences of thematic R&D programs. Even with the same technology field firms apply and agency select R&D projects with higher private value.

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, our analysis suffers from the fact that we cannot disentangle the selection effect and reverse causality due to data limitations. Nevertheless, we hope our analysis serves as inspiration for future research in the field of analysing the effects of R&D subsidies on the project level. The question which differences between subsidized and non-subsidized R&D projects are due to selection choices of firms and funding agencies or due to the effect of the subsidy on the project outcome calls for further research and additional project-level data for both subsidized and non-subsidized R&D projects. In our view this is necessary in order to provide a more tailored policy advice to our governments when it comes to support private R&D in our economies best.

References

Aerts, K., and T. Schmidt (2006). Two for the price of one? Additionality effects of R&D subsidies: a comparison between Flanders and Germany. *Research Policy* 37(5), 806–822.

Almus, M., and D. Czarnitzki (2003). The effects of public R&D subsidies on firm's innovation activities. The case of Eastern Germany. *Journal of Business and Economic Statistics* 21(2), 226–236.

Arqué-Castells, P., and P. Mohnen (2015). Sunk costs, extensive R&D subsidies and permanent inducement effects. *Journal of Industrial Economics* 63(3), 458–494.

Arrow, K. (1962). Economic welfare and the allocation of resources to invention. In: R. Nelson (ed.), *The rate and direction of incentive activity*. Princeton: Princeton University Press.

Aschhoff, B. (2009). *The effects of subsidies on R&D investments and innovation success – Do subsidy history and size matter?* ZEW Discussion Paper 09-032, Mannheim.

Aschhoff, B. (2010). Who gets the money? The dynamics of R&D project subsidies in Germany. *Jahrbücher für Nationalökonomie und Statistik* 230(5), 522-546.

Aschhoff, B., M. Astor, D. Crass, T. Eckert, S. Heinrich, G. Licht, C. Rammer, D. Riesenberg, N. Rueffer, R. Strohmeyer, V. Tonoyan, and M. Woywode (2012). *Systemevaluierung „KMU-innovativ"*. ZEW Documentation 12-04.

BMBF (2010). *Bundesbericht Forschung und Innovation 2010*. BMBF: Berlin.

BMBF (2012). *Bundesbericht Forschung und Innovation 2012*. BMBF: Berlin.

Branstetter, L., and M. Sakakibara (1998). Japanese research consortia a microeconometric analysis of industrial policy. *The Journal of Industrial Economics* 46(2), 207-233.

Branstetter, L., and M. Sakakibara (2002). When do research consortia work well and why? Evidence from Japanese panel data. *American Economic Review* 92(1), 143-159.

Czarnitzki, D., and M. Fier (2002). Do innovation subsidies crowd out private investment? Evidence from the German service sector. *Applied Economics Quarterly* 48(1), 1–25.

Czarnitzki, D., and K. Hussinger (2017). Input and output additionality of R&D subsidies. *Applied Economics* published online 16 August 2017, 1-18.

Czarnitzki, D., and G. Licht (2006). Additionality of public R&D grants in a transition economy: The case of Eastern Germany. *Economics of Transition* 14(1), 101-131.

Czarnitzki, D., B. Ebersberger, and A. Fier (2007). The relationship between R&D collaboration, subsidies and R&D performance: empirical evidence from Finland and Germany. *Journal of Applied Econometrics* 22(7), 1347–1366.

Czarnitzki, D., and C. Lopes-Bento (2014). Innovation subsidies: Does the funding source matter for innovation intensity and performance? Empirical evidence from Germany. *Industry and Innovation* 21(5), 380-409.

Cohen, W. M., and R. C. Levin (1989). Empirical studies of innovation and market structure. In: *Handbook of Industrial Organization*, Vol.2. Edited by: R. Schmalensee and R.D. Willig, Elsevier Science Publishers, 1059-1107.

David, P. A., B. H. Hall, and A. A. Toole (2000). Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Research Policy* 29(4-5), 497-529.

EC (1986). *Mitteilung der Kommission – Gemeinschaftsrahmen für staatliche Forschungs- und Entwicklungsbeihilfen*. Official Journal of the European Union, C83 (11 April 1986), 2-5.

EC (1993a). *Council Regulation (EEC) No 2083/93 of 20 July 1993 amending Regulation (EEC) No 4254/88 laying down provisions for implementing Regulation (EEC) No 2052/898 as regards the European Regional Development Fund*. Official Journal of the European Communities, L193 (31 July 1993), 34-38.

EC (1993b). *Council Regulation (EEC) No 2084/93 of 20 July 1993 amending Regulation (EEC) No 4255/88 laying down provisions for implementing Regulation (EEC) No 2052/88 as regards the European Social Fund*. Official Journal of the European Communities, L193 (31 July 1993), 39-43.

EC (1996). *Mitteilung der Kommission – Gemeinschaftsrahmen für staatliche Forschungs- und Entwicklungsbeihilfen*. Official Journal of the European Union, C045 (17 February 1996), 5-16.

EC (1999a). *Regulation (EC) No 1783/1999 of the European Parliament and of the Council of 12 July 1999 on the European Regional Development Fund*. Official Journal of the European Communities, L213 (13 August 1999), 1-4.

EC (1999b). *Regulation (EC) No 1784/1999 of the European Parliament and of the Council of 12 July 1999 on the European Social Fund*. Official Journal of the European Communities, L213 (13 August 1999), 5-8.

EC (2006). *Mitteilung der Kommission – Gemeinschaftsrahmen für staatliche Beihilfen für Forschung, Entwicklung und Innovation*. Official Journal of the European Union, C 323 (30 December 2006), 1-26.

Eurostat (2017). *Total intramural R&D expenditure (GERD) by sectors of performance [rd_e_gerdtot].* http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=rd_e_gerdtot&lang=en from 5 October 2017.

Fier, A. (2002). *Staatliche Förderung industrieller Forschung in Deutschland: Eine empirische Wirkungsanalyse der direkten Projektförderung des Bundes*. Nomos Verlagsgesellschaft: Baden-Baden.

Giuri, P., and M. Mariani (2005). *Everything you always wanted to know about inventors (but never asked): Evidence from the PatVal-EU Survey*. LEM Working Paper Series 2005/20.

Goolsbee, A. (1998). Does Government R&D Policy Mainly Benefit Scientists and Engineers? *American Economic Revie*w 88(2), 298-302.

Hagedoorn, J. (2003). Sharing intellectual property rights—an exploratory study of joint patenting amongst companies. *Industrial and Corporate Change* 12(5), 1035-1050.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). Market value and patent citations. *The RAND Journal of Economics* 36(1), 16-38.

Hud, M., and K. Hussinger (2015). The impact of R&D subsidies during the crisis. *Research Policy* 44(10), 1844-1855.

Hussinger, K. (2008). R&D and subsidies at the firm Level: An application of parametric and semiparametric two-step selection models. *Journal of Applied Econometrics* 23(6), 729-747.

Jaffe, A., M. Trajtenberg, and R. Henderson (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics* 108(3), 577-598.

Jaffe, A., M. S. Fogarty, and M. Trajtenberg (2000). Knowledge spillovers and patent citations: evidence from a survey of inventors. *American Economic Review* 90(2), 215-118.

Klette, T. J., J. Moen, and Z. Griliches (2000). Do subsidies to commercial R&D reduce market failures? Microeconometric evaluation studies. *Research Policy* 29(4-5), 471-495.

Lanjouw, J. O., and M. Schankerman (1999). *The quality of ideas: measuring innovation with multiple indicators*. NBER Working Paper W7345.

Mansfield, E., M. Schwarz, and S. Wagner (1981). Imitation costs and patents: An empirical study. *The Economic Journal* 91(December), 907-918.

Mansfield, E. (1985). How rapidly does new industrial technology leak out?. *The Journal of Industrial Economics* 34(2), 217-223.

Nagaoka, S., K. Motohashi, and A. Goto (2010). Patent statistics as an innovation indicator, in: *Handboock of Economics of Innovation*. Edt: B.H. Hall and N. Rosenberg. Volume 2, chapter 25, p. 1083-1127.

NKFT88 (1988). *Nebenbestimmungen für Zuwendungen auf Kostenbasis des Bundesministeriums für Forschung und Technologie an Unternehmen der gewerblichen Wirtschaft für Forschungs- und Entwicklungsvorhaben*. BMBF/BMFT-Vordruck 0348/01.89.

NKBF98 (1998). *Nebenbestimmungen für Zuwendungen auf Kostenbasis des Bundesministeriums für Forschung und Technologie an Unternehmen der gewerblichen Wirtschaft für Forschungs- und Entwicklungsvorhaben*. BMBF-Vordruck 0348a/06.06_2.

PatVal-EU (2005). *The value of European patents: Evidence from a survey of European inventors*. Final Report of the PatVal EU Project. http://ec.europa.eu/invest-in-research/pdf/download_en/patval_mainreportandannexes.pdf.

Schmoch, U. (2008). *Concept of a technology classification for cross country comparisons*. Final report to the World Intellectual Property Organization (WIPO). Available at WIPO: http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/pdf/wipo_ipc_technology.pdf.

Schneider, C. (2008). *Mixed R&D incentives: The effect of R&D subsidies on patented inventions*, (October 14, 2008). Available at SSRN: http://ssrn.com/abstract=1311784.

Takalo, T., T. Tanayama, and O. Toivanen (2013a). Estimating the benefits of targeted R&D subsidies. *Review of Economics and Statistics* 95(1), 255-272.

Takalo, T., T. Tanayama, and O. Toivanen (2013b). Market failures and the additionality effects of public support to private R&D: Theory and empirical implications. *International Journal of Industrial Organziation* 31, 634-642.

Trajtenberg, M., R. Henderson, and A. Jaffe (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology* 5, 19-50.

Veugelers, R. (1998). Collaboration in R&D: An assessment of theoretical and empirical findings. *De Economist* 146(3), 419-443.

Zúñiga-Vicente, J. Á., C. Alonso-Borrego, F. J. Foracadell, and J. I. Galán (2014). Assessing the effect of public subsidies on firm R&D investment: A survey. *Journal of Economic Surveys* 28(1), 36-67.

Appendix 1

Table 4: Descriptive statistics of patent applications from subsidized and non-subsidized R&D projects 1995-2005 (sample 2)

Variable	All				Subsidized				Non-Subsidized				t-
	mean	s.d.	min	max	mean	s.d.	min	max	mean	s.d.	min	max	Test
Private value													
SELFCIT	0.309	0.912	0	31	0.367	0.995	0	9	0.308	0.910	0	34	*
TRIADIC	0.280	0.449	0	1	0.304	0.460	0	1	0.280	0.449	0	1	
Spillovers													
EXTCIT	0.599	1.348	0	45	0.579	1.204	0	13	0.599	1.350	0	45	
COAPP	0.023	0.149	0	1	**0.071**	0.257	0	1	0.022	0.146	0	1	***
Basicness													
GENERAL	0.044	0.136	0	0.813	0.050	0.146	0	0.778	0.044	0.136	0	0.813	
NOFORW	0.587	0.492	0	1	0.574	0.495	0	1	0.587	0.492	0	1	
ORIGINAL	0.109	0.208	0	0.953	**0.146**	0.237	0	0.878	0.109	0.207	0	0.953	***
BACKCIT	1.815	2.569	0	36	**2.352**	2.944	0	28	1.805	2.560	0	36	***
NOBACK	0.515	0.500	0	1	0.398	0.490	0	1	**0.517**	0.500	0	1	***
Project size													
INVCNT	2.641	1.820	1	21	**3.008**	1.852	1	14	2.634	1.819	1	21	***
No. of obs.	61,065				1,127				59,938				
No. of firms	206				206				206				

Notes: t-test displays the level of significance of a mean difference test between the group of patent applications from subsidized and non-subsidized projects. *p<0.10, **p<0.05, ***p<0.01. If the difference between patents from subsidized and non-subsidized R&D projects is significant above the 5%-level, we bold printed the higher of the two means.

Table 5: Proportion of patent applications from non-subsidized and subsidized R&D projects 1995-2005, by technology fields

Variable		All		Subsidized	Non-Subsidized	t-test
		mean	share of subsidized patents	mean	mean	
Electrical engineering						
1	Electrical machinery, apparatus and energy	0.097	1.3%	0.181	0.097	***
2	Audio-visual technology	0.017	0.3%	0.008	0.017	*
3	Telecommunications	0.084	0.8%	0.096	0.084	
4	Information technology	0.046	0.6%	0.041	0.046	
5	Semiconductors	0.046	2.0%	0.133	0.045	***
Instruments						
6	Optics	0.026	1.9%	0.070	0.026	***
7	Analysis, measurement, control technology	0.103	0.8%	0.120	0.102	*
8	Medical technology	0.016	0.2%	0.004	0.016	***
9	Nuclear engineering	0.006	1.1%	0.009	0.006	
Chemistry & Pharmaceutical						
10	Organic fine chemistry	0.050	0.4%	0.029	0.050	***
11	Macromolecular chemistry, polymers	0.050	0.4%	0.031	0.050	**
12	Pharmaceuticals, cosmetics	0.024	0.3%	0.012	0.024	**
13	Biotechnology	0.010	2.7%	0.040	0.010	***
14	Agriculture, food chemistry	0.002	1.3%	0.003	0.002	
15	Chemical and petrol industry, basis	0.028	0.7%	0.029	0.028	
16	Surface technology, coating	0.017	1.6%	0.040	0.017	***
17	Materials, metallurgy	0.016	1.9%	0.045	0.016	***
Process engineering & special equipment						
18	Chemical engineering	0.030	1.0%	0.043	0.030	**
19	Materials processing, textiles, paper	0.034	0.4%	0.020	0.034	**
20	Handling, printing	0.021	0.3%	0.009	0.021	**
21	Agricultural and food processing, apparatus	0.001	0.7%	0.001	0.001	
22	Environmental technology	0.013	0.8%	0.014	0.013	
Mechanical engineering & machinery						
23	Machine tools	0.021	0.4%	0.012	0.021	*
24	Engines, pumps, turbines	0.102	0.2%	0.032	0.102	***
25	Thermal processes and apparatus	0.019	1.8%	0.050	0.019	***
26	Mechanical elements	0.058	0.2%	0.019	0.058	***
27	Transport	0.164	0.1%	0.015	0.165	***
28	Space technology, weapons	0.002	0.3%	0.001	0.002	
Consumption						
29	Consumer goods and equipment	0.011	1.4%	0.021	0.010	***
30	Civil engineering, building, mining	0.012	0.9%	0.016	0.012	
No. of obs.		174.311		1.207	173.104	
No. of firms		249		249	249	

Notes: Technological areas according to the classification suggested by Schmoch (2008). All values are calculated using sample 1. A patent application may be assigned to more than one technological area. t-test denotes a mean difference test between the group of patent applications from subsidized and non-subsidized projects. *p<0.10, **p<0.05, ***p<0.01.

Table 6: Robustness analysis for sample 1

	(1) Probit	(2) OLS	(3) Probit		(4) Probit	(5) Probit
	Whole sample 1 (see Table 4)	Whole sample 1	Split between 10% highest and lower patenting firms		Excluding patents from control group that cite a subsidized patent	Subsample 2000-2005
			low	high		
Private value						
SELFCIT	0.016	0.033*	0.125	0.000	0.016	0.034
	(0.020)	(0.020)	(0.068)	(0.021)	(0.020)	(0.030)
TRIADIC	0.128**	0.121**	0.066	0.123**	0.129**	0.208***
	(0.051)	(0.048)	(0.191)	(0.048)	(0.051)	(0.078)
Spillovers						
EXTCIT	-0.017	-0.021*	-0.071	-0.011	-0.016	-0.035
	(0.018)	(0.011)	(0.060)	(0.019)	(0.018)	(0.036)
COAPP	0.413***	0.568***	1.228***	0.140	0.420***	0.242
	(0.129)	(0.154)	(0.405)	(0.123)	(0.129)	(0.160)
Basicness						
GENERAL	0.345**	0.379**	1.355**	0.157	0.347**	0.594**
	(0.159)	(0.163)	(0.579)	(0.158)	(0.160)	(0.245)
NOFORW	-0.066	-0.069	-0.199	-0.049	-0.065	-0.049
	(0.049)	(0.046)	(0.179)	(0.048)	(0.049)	(0.074)
ORIGINAL	-0.006	-0.023	-0.006	-0.002	-0.011	0.047
	(0.107)	(0.136)	(0.402)	(0.102)	(0.107)	(0.145)
BACKCIT	0.002	0.000	0.053	-0.011	0.004	-0.008
	(0.010)	(0.012)	(0.034)	(0.011)	(0.010)	(0.014)
NOBACK	-0.271***	-0.302***	-0.435**	-0.243***	-0.276***	-0.358***
	(0.058)	(0.069)	(0.222)	(0.056)	(0.058)	(0.080)
Project size						
INVCNT	0.075***	0.076***	0.288***	0.040***	0.075***	0.096***
	(0.011)	(0.011)	(0.043)	(0.010)	(0.011)	(0.015)
Year dummies	yes	yes	yes	yes	yes	yes
Area dummies	yes	yes	yes	yes	yes	yes
Firm dummies	yes	yes	yes	yes	yes	yes
McFadden's R^2	0.230		0.293	0.144	0.231	0.270
McK & Zav R^2	0.261		0.343	0.245	0.262	0.283
Count R^2	77.11%		80.74%	72.98%	77.23%	77.61%
R^2		0.067				
No. of obs.	174,311	174,311	30,160	139,190	173,961	100,873
No. of firms	249	249	224	25	248	182

Notes: Displayed are average marginal effects and standard errors for Probit and coefficients and robust standard errors for OLS estimation. Marginal effects, coefficients and standard errors are multiplied by 100 for ease of representation. *p<0.10, **p<0.05, ***p<0.01. See also notes for Table 3.

Table 7: Robustness analysis for sample 2

	(1) Probit	(2) OLS	(3) Probit		(4) Probit	(5) Probit	(6) Probit
	Whole sample 2 (see Table 4)	Whole sample 2	Split between 10% highest and lower patenting firms		Excluding patents from control group that cite a subsidized patent	Subsample 2000-2005	Random draw of max. 5 control patents per subsidized patent
			low	high			
Private value							
SELFCIT	0.015	0.039	0.407	-0.006	0.015	0.050	0.219
	(0.057)	(0.063)	(0.421)	(0.051)	(0.057)	(0.070)	(0.259)
TRIADIC	0.462***	0.421***	-0.225	0.414***	0.467***	0.646***	4.149***
	(0.142)	(0.134)	(0.851)	(0.129)	(0.142)	(0.174)	(0.581)
Spillovers							
EXTCIT	-0.024	-0.041	-0.264	-0.005	-0.022	-0.159*	-0.180
	(0.051)	(0.043)	(0.361)	(0.044)	(0.051)	(0.083)	(0.248)
COAPP	0.895**	1.431***	5.816***	0.333	0.930**	0.277	7.816***
	(0.369)	(0.536)	(1.936)	(0.326)	(0.374)	(0.344)	(1.451)
Basicness							
GENERAL	0.712	0.723	1.583	0.658	0.716	1.456***	5.890***
	(0.447)	(0.470)	(2.781)	(0.402)	(0.448)	(0.537)	(1.830)
NOFORW	-0.092	-0.093	-0.472	-0.086	-0.090	-0.067	-0.749
	(0.131)	(0.126)	(0.846)	(0.119)	(0.132)	(0.160)	(0.558)
ORIGINAL	0.078	0.131	1.729	-0.150	0.052	0.129	0.066
	(0.291)	(0.348)	(1.671)	(0.273)	(0.292)	(0.306)	(1.165)
BACKCIT	-0.035	-0.029	0.098	-0.052*	-0.028	-0.043	-0.221**
	(0.029)	(0.036)	(0.143)	(0.029)	(0.029)	(0.031)	(0.108)
NOBACK	-0.770***	-0.780***	-2.576***	-0.6186***	-0.775***	-0.770***	-7.014***
	(0.159)	(0.173)	(0.974)	(0.147)	(0.160)	(0.167)	(0.590)
Project size							
INVCNT	0.188***	0.202***	0.954***	0.117***	0.188***	0.191***	1.949***
	(0.030)	(0.031)	(0.190)	(0.027)	(0.030)	(0.032)	(0.155)
Year dummies	yes	yes	yes	yes	yes	yes	yes
Area dummies	yes	yes	yes	yes	yes	yes	yes
Firm dummies	yes	yes	yes	yes	yes	yes	yes
McFadden's R^2	0.197		0.335	0.085	0.198	0.237	0.090
McK & Zav R^2	0.221		0.52	0.145	0.222	0.255	0.197
Count R^2	75.64%		79.07%	65.16%	75.60%	76.37%	64.25%
R^2		0.102					
No. of obs.	61,065	61,065	6,000	54,775	60,900	46,161	6,018
No. of firms	206	206	185	21	206	159	206

Notes: Displayed are average marginal effects and standard errors for Probit and coefficients and robust standard errors for OLS estimation. Marginal effects, coefficients and standard errors are multiplied by 100 for ease of representation. See also notes for Table 3. *p<0.10, **p<0.05, ***p<0.01. Column 6 displays the average results of 200 iterations.

Appendix 2

As discussed in section 6.3 our data suffers from potential measurement error because we may not be able to identify all patents that stem from subsidized patents. In order to understand how this may affect our results, we quantify the size of the potential measurement error in our dependent variable SUBS.

For this purpose we use information from the PatVal survey (PatVal-EU 2005). In the PatVal survey inventors of 3,346 patents with priority date in the years 1993 to 1997 located in Germany were asked about the financing sources of the project that led to the filing of the patent. 5.41% of German inventors answered that their patent resulted from Government Research Programs or other government funds. At the same time 3.51% of the German patents are filed by inventors employed at a public research institute. Assuming that these patents stem from public funding, we find that a share of 1.9% of the German patents are from publicly funded projects of inventors not employed in public research institutes.

Recall that the PatVal sample differs from our sample. PatVal is a *random* sample of patents in Germany. In contrast, our sample contains only patents of firms that have filed *at least one patent from a subsidized project*. Hence, the share of subsidized patents in our sample should be larger than 1.9%. Based upon the PatVal information we try to estimate the share of subsidized patents for a sample of patents from firms that filed at least one subsidized patent.

First, we create a random sample of patenting firms in Germany. In order to do so we use the Mannheim Innovation Panel (MIP), a random sample of German firms, between 1992 and 2005 and merge this dataset with all patents filed at the GPO at that time. The final sample contains 343,769 patent applications filed by 6,358 firms. Assuming that the PatVal survey figure is correct and constant across time, we can calculate the number of patents (*subpat*) that should actually be subsidized in the sample of patenting MIP firms according to the equation $\widehat{subpat} = 0.019 * allpat$.

allpat is the number of all patent applications at the GPO from firms in the sample between 1992 and 2005. Considering all 343,769 patent applications from 6,358 firms filed between 1992 and 2005, we *should* observe 6,532 subsidized patents (1.9%). In contrast, we can only identify 1,240 subsidized patents filed by 196 firms (0.36%).

Second, in order to estimate the potential measurement error for our sample, we need to estimate the share of subsidized patents for the subsample of *those firms that filed at least one subsidized patent*. Which firms should have filed the 6,532 subsidized patents? At minimum, it should be those 196 firms that filed the 1,240 subsidized patents, which we are able to identify. At maximum, it should be all those firms, which indicated in the MIP questionnaire that they have received public R&D funding from the EU, Federal or the German States in the period of observation (1,197 firms). Hence, patents filed by *firms that received an R&D subsidy* are: At the minimum 195,329 patents filed by 196 firms, and at the maximum 245,842 patents filed by 1,197 firms.

This allows us to estimate a minimum and a maximum bound of the true share of subsidized patents in a sample of patents filed by firms that received an R&D subsidy. The estimated share of subsidized patents compared to patents filed by firms that received an R&D subsidy is minimum 2.7% and maximum 3.3%. In comparison, the share of subsidized patents in sample 1 from 1995-2005 is 0.7%. Hence, we conclude that the measurement error in our subsidy variable (SUBS) ranges between 2.0-2.6%.

Appendix 3

Correlation of indicators

The employed patent indicators are partly based on the same sources. For instance, BACKCIT and ORIGINAL are both based on backward citations while GENERAL, SELFCIT and EXTCIT make use of forward citations. Table 8 therefore depicts the pairwise correlation between the patent indicators for sample 1. Correlations in sample 2 are comparable to those of 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 8: Correlation matrix of patent indicators

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

Notes: All values are calculated using sample 1. *p<0.05.