The Behavioral Additionality of Government Research Grants[*]

Rainer Widmann[‡]

Max Planck Institute for Innovation and Competition

January 2023

Abstract

There are different forms of public support for industrial R&D. Some attempt to increase innovation by prompting firms to undertake more challenging projects than they would otherwise do. Access to a dataset from one such program, the Austrian Research Promotion Agency, allows me to examine the effect of research grants on firms' patenting outcomes. My estimates suggest that a government research grant increases the propensity to file a patent application with the European Patent Office by around 12 percentage points. Stronger effects appear for more experienced firms of advanced age. Additional evidence indicates that grants induce experienced firms to develop unconventional patents and patents that draw on knowledge novel to the firm. I interpret the findings in a "exploration vs. exploitation" model, in which grants are targeted at ambitious projects that face internal competition from more conventional projects within firms. The model shows that this mechanism is more salient in experienced firms, leading to a stronger response in behavior for this group of firms.

[*]I thank Benjamin F. Jones, Craig Garthwaite, Nicola Persico, David Besanko, Dietmar Harhoff, Matthias Doepke, Pere Arqué-Castells, Nicola Bianchi, Matthew Notowidigdo, Harald Oberhofer, Fabian Gässler, Emilio Raiteri and participants of the 8th ZEW/MaCCI conference, the Kellogg strategy brown bag seminar, the VPDE conference Torino, the MPI Innovation Seminar, the Mannheim Summer School on Taxation and R&D and the research seminar at the Austrian Institute of Economic Research (WIFO). I am grateful to the staff of the Austrian Research Promotion Agency (FFG), especially Josef Säckl and the strategy unit, for their assistance and hospitality. Financial support was provided by the Kellogg strategy department and the General Motors Research Center. All errors are mine.

[†]Email: rainer.widmann@ip.mpg.de

[‡]Declaration of interest: I declare that I have no competing financial interest or personal relationship that could have influenced the work reported in this paper.

1 Introduction

Since firms cannot appropriate all gains generated by inventions, private investment in Research & Development falls short of the socially optimal level of investment (see, for example, Bloom et al. 2013, Hall 2002, Jones and Williams 1998, Jaffe 1986, Nelson 1959). For this reason, all developed countries have put policies in place that support and enable firm innovation. Governments regularly subsidize R&D in order to raise the rate of innovation. Subsidies are provided directly through grants, allocated by government agencies, and indirectly through tax credits. In this paper, I study the effects of a particular directed government research grant program.

Policy makers around the world ceaselessly experiment with new funding schemes, bringing about considerable heterogeneity in research grant programs. Studies have documented the effects of these programs (see Section 2), but there are still policy-relevant mechanisms that are not well understood. Recent policy initiatives have targeted young and small firms, mainly based on the view that R&D subsidies are most impactful when aimed at financially constrained firms. The corresponding programs typically offer grants that cover the cost of developing a prototype or an invention, and do not require co-financing. Studies by Howell (2017) of the SBIR program in the US and Santoleri et al. (2021) of the SME instrument of the European Commission have confirmed the positive effect of these programs on innovation output, suggesting that they are effective in alleviating financial frictions.

At the same time, numerous European countries have long-standing industrial R&D support programs that are decidedly broader in scope, funding research also in experienced and large firms. National R&D subsidy programs in the European Economic Area must comply with the EU directives on state aid, which prescribe substantial co-financing on the part of the subsidized firm.[1] In many cases, this creates a tension for the program between selecting projects for which private sector co-financing is available, and selecting projects which would not be undertaken in the absence of funding. Given this limitation, it appears unlikely that the alleviation of financial frictions is the main mechanism through which these programs affect innovation.

An established line of reasoning suggests that, even in the absence of financial frictions, public funding may enable additional projects which are of low value to firms, but which entail spillovers (Jaffee 1998). Another possibility is that programs prompt firms to undertake more challenging or ambitious projects than they would otherwise do, thereby increasing the chances that they develop inventions that are of high value (to them). This requires that firms manage portfolios of projects of varying degree of ambition, each directly and indirectly affected by public funding. The view, that public funding may lead to an increase in scale or scope of projects is an important aspect of the concept of "behavioral additionality", introduced by Buisseret, Cameron & Georghiou (1995).

[1] Official Journal of the European Communities, C 45/5, 2/1996: Community framework for state aid for research and development 96/C45/06. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31996Y0217(01)&qid=1623663125993&from=EN

I leverage unique data from the Austrian Research Promotion Agency (FFG) to shed light on this issue. The institution studied in this paper dispenses grants to a broad population of firms, ranging from young startups to older, more experienced firms. The program is meant to encourage high-risk projects that "push the envelope beyond [the firms'] normal R&D activities" (Arnold et al. 2004, p. 54). In well-resourced firms, the agency stirs the selection of projects towards more challenging and unfamiliar terrain. The grants entail considerable co-financing, as they fund only 25% of project cost, but scale to the size of projects without placing an upper limit on cost. Their design is similar to industrial R&D support programs of other European national agencies.[2]

I analyze 2619 funding applications across 1936 firms for public research grants between 2002 and 2005. To identify the effect of the grant, I exploit a steep, almost discontinuous increase in the dependence of funding approval on the agency's internal evaluation score. I find that research grants have an effect of around 12 percentage points on the propensity to file a patent with the European Patent Office (EPO) within 4 years. In the Austrian context, applications at the European level indicate that inventions are of high value to firms. The effect is particularly pronounced for firms that are above the median age in the sample, which is 5 years. On the other hand, estimates of the effect for younger firms appear economically small and are statistically insignificant.

Additional evidence appears broadly consistent with claims that publicly funded projects are technically more challenging and "allow [firms] to expand their R&D into new fields" (Arnold et al. 2004, p. 53). There is a discernible effect on the firms' propensity to file patents that are "unconventional", as measured by how atypical combinations of cited technological classes that appear together on the patent. On the other hand, there appears to be no increase in the propensity to file patents that are relatively conventional. Although not entirely conclusive, firms also appear more likely to cite technological classes that are "novel to the firm", i.e. classes that they had never cited prior to the grant.

I interpret the findings in a model in the vein of Buisseret, Cameron & Georghiou (1995), in which ambitious projects face internal competition from more conventional projects within firms. In the model, firms incessantly choose between undertaking technically related, incremental projects on their existent research line, which are safe ("exploitation"), and undertaking risky projects that potentially lead to the discovery of a new research line ("exploration"). Because new research lines replace old research lines, the incentive of experienced firms to "innovate again" is diminished, akin to Arrow's replacement effect (Arrow 1962). This hurdle is addressed by matching grants that target risky projects. The policy can be rationalized by assuming that the social value of research declines as research lines become more outdated and projects become more

[2] The Finish agency Tekes covers 25-50% of the cost of industrial R&D projects. Projects that involve technological risks are specifically supported (Einiö 2014). The Belgian agency IWT ("Agency for Innovation by Science and Technology in Flanders") funds up to 35% of cost. The evaluation criteria focus on scientific qualities and technological risks of projects, aside from commercial criteria (Czarnitzki and Lopes-Bento 2013). Other European agencies that pay matching grants to young and experienced firms for industrial R&D projects are Vinnova (Sweden) and the Research Council of Norway. A list of national government agencies in the European Union that fund research, in some cases also in the academic sector, can be found on the homepage of the Research and Innovation Observatory of the EU Commission (URL: `https://rio.jrc.ec.europa.eu`, accessed on 6/30/2021).

and more incremental in nature.

In addition, I use the model to derive predictions on the effectiveness of grants when the agency is constrained to a fixed matching rate and when there is imperfect compliance. The model predicts that grants to young firms may be ineffective in the initial years after their incorporation. For young firms, the value of exploiting their recently established research line is high when compared to the more dated research lines of older firms. A fixed-rate grant is therefore less likely to provide sufficient incentive to shifting to a new research line for young firms. Young firms may also lack competing conventional projects, rendering grants superfluous. The model predicts that, as firm cohorts age, there is a strictly positive share of firms which behavior is affected by the grant, and consequently a strictly positive effect of grants on patenting outcomes.

The rest of the paper is organized as follows. In section 2, I discuss related literature. Section 3 describes the Austrian Research Promotion Agency FFG and the data. In Section 4, I lay out the empirical strategy. Section 5 contains the results of the effect of research grants. Section 6 presents the model. Section 7 concludes.

2 Related literature

This paper contributes to the literature evaluating directed R&D subsidy programs. A large body of evidence examines whether subsidies increase R&D expenditures.[3] However, when comparing funded and unfunded firms, the selection confounds the causal effect that can be attributed to the subsidy program. For this reason, recent studies focus on designs in which institutional features generate exogenous variation in funding. Bronzini and Iachini (2014) examine a research grant program in northern Italy, exploiting an evaluation score cutoff, and find no significant increase in R&D expenditures overall. Utilizing exogenous variation in subsidies generated by population-density rules, Einiö (2014) finds evidence of a positive effect of R&D subsidies on R&D expenditures for the finish agency Tekes.

In comparison, the empirical literature on the effect of R&D subsidies on innovation output is less voluminous. An important contribution is Howell (2017), who studies the US SBIR program, also using a Regression Discontinuity design. The program is aimed at small, innovative firms and offers large, fixed size grants that entail no co-financing. The US SBIR has served as a model for policy initiative elsewhere. The most notable example is the EU SME program, which is analyzed in Santoleri et al. (2021). Both studies find a positive effect on patent outcomes and show that young recipient firms are subsequently more likely to receive outside venture capital. Bronzini and Piselli (2016) find that a research program in northern

[3]Zúñiga-Vicente et al. (2014) survey 76 descriptive studies and conclude that evidence on the effect is inconclusive. The authors suppose that differences in the designs of the subsidy programs and differences in the studied firm populations account for the heterogeneity in the findings Important studies included in the survey are Wallsten (2000) and Lach (2002) who do not find an effect of subsidies on R&D expenditures, and Lichtenberg (1988), Almus and Czarnitzki (2003), González et al. (2005), Hussinger (2008) who find positive and statistically significant effects. For a structural model of subsidies, see Takalo, Tanayama and Toivanen (2013a).

Italy increased the number of patent applications and the propensity to patent.[4] They find stronger effects for smaller firms, which they attribute to financial frictions faced by small firms. Additional studies on the effect of directed subsidies on other measures of innovative activity include Lerner (1999), Demeulemeester and Hottenrott (2017) and Hünermund and Czarnitzki (2019).

The present paper, especially the theoretical part, draws on ideas presented in the foundational contribution of Buisseret, Cameron & Georghiou (1995) on "behavioral additionality", and to the related report by Georghiou (2002). Buisseret, Cameron & Georghiou (1995) define the concept as "the change in a company's way of undertaking R&D which can be attributed to policy actions" (p. 590). Changes in firm behavior are conceived as extending beyond the implementation of a singular, isolated project.[5] They consider a project portfolio setting where subsidies directly or indirectly affect supported projects and other R&D activity within the recipient firm. In their framework, public funding can lead to a reprioritisation towards projects that exhibit different properties. In Georghiou (2002), the author argues that the "common effect of innovation policy [is] not to alter a stop-go decision by the firm in respect of the project" (p.59), but to increase their scale, their scope or the speed with which they are undertaken. Since then, others have examined how public funding can affect firm behavior via changes in organizational practices (Falk 2007, Neicu, Teirlinck & Kelchtermans 2016) or via learning (Clarysse, Wright & Mustar 2009).

The findings in this paper complement results from the literature on R&D tax credits. There is an extensive literature assessing the impact of R&D tax credits on R&D expenditures, both on the macroeconomic level and on the firm level (surveyed in Becker 2014). Bloom et al. (2002) find evidence that tax credits increased aggregate R&D expenditures in a sample of OECD countries between 1979 and 1997. Rao (2016) finds large positive elasticities of firms' R&D expenditures with respect to tax credits, exploiting changes in federal tax advantages in the US between 1981 and 1991. Studies that examine the effect of R&D tax credits on innovation output include Dechezleprêtre et al. (2016), Capellen et al. (2012) and Czarnitzki et al. (2011).

3 Institutional setting and data

3.1 The Austrian Research Promotion Agency FFG

The Austrian Research Promotion Agency FFG ("Forschungsförderungsgesellschaft") plays a dominant role in the public funding landscape for R&D in Austria, accounting for 80% of all direct government funding for Business R&D (Arnold et al. 2004). The main vehicle of funding is the "Basisprogramm". The "Ba-

[4]The institutional context and the set of firms analyzed in Bronzini and Iachini (2014) and Bronzini and Piselli (2016) appear to share some of the features of the institution and the characteristica of firms studied in this paper, but differ in some respects: First, the median age of the firms included in their sample is around 16 years. Second, the upper limit on eligible costs for one project in their program is only 250,000 Euro and projects are funded at a relatively high matching rate of 50%.

[5]Most papers study the implementation of a singular project in the presence of spillovers and financial frictions. Takalo et al. (2013b) study the problem of funding a project in a model that incorporates spillovers and financial frictions. Besanko et al. (2018) compare subsidy policies in an exponential bandit model where a decision maker decides when to give up on a project that has spillovers. Lach et al. (2017) analyze government financing of an R&D project when the entrepreneur is privately informed about the riskiness of her project. An exception is Bryan & Lemus (2017), who study innovation policy in a model where the firm choses between multiple projects (correspondent to different "directions of research") .

sisprogramm" is a project-oriented research grant program with rolling admission, open to all firms of all sectors, that has existed in a relatively consistent form since the early 1990s. Firms apply for funding with proposals for research projects with no formal constraints on research area or topic. All subsequent analysis is carried out only for applications to the "Basisprogramm".

According to the report of Arnold et al. (2004), the program aims to reward high-risk projects in both smaller and more established companies: "The way [that the agency] tailors its funding instruments complies with the normal funding rationale: [it] rewards high risk" (Arnold et al. 2004, p. 37). Unlike programs that focus exclusively on start-ups, "[the agency] devote[s] significant subsidy resources to the high-potential internal projects of large and well-resourced companies" (Arnold et al. 2004, p. 50). The authors claim that grants enable technically more difficult projects and allow firms to extend their R&D into new fields. Accordingly, grants help firms "push the envelope beyond their normal R&D activities" (Arnold et al. 2004, p. 54). The analysis in this paper examines if the stated perceptions are in line with actual outcomes.

Funding applications include a detailed description of the project and a quote of the total cost. The agency applies a standardized evaluation procedure to all applications, which reflects the objectives of the agency. The technical assessment of the project proposal is based on the categories "risk/challenge", "novelty", "practical value" and "environmental effects". This assessment, which my empirical strategy is based on, is discussed in detail further below. In a separate assessment, the proposed project is also evaluated relative the firm's other R&D efforts, the most important aspect being the "increase in know-how" for the firm. Aside from the technical assessment, there is a commercial assessment, which is meant to ensure that the firm is able to finance the remaining share of the cost of the proposed project.

The technical assessment of the project proposal results in an evaluation score called $pwert$. The first step in the evaluation is that the application is assigned to the responsible examiner. The agency has in-house expert examiners for each technological field. The examiner then works through a list of pre-defined technical characteristics. Each technical characteristic is awarded a value between 0 and 4. The assessments of the separate technical characteristics are aggregated into the four categories listed above, where each category comprises between 2 and 5 characteristics. The category scores are placed on discrete grids of admissible integer values. The $pwert$ score is then given by the sum of the category scores. The $pwert$ score ranges between 0 and 50. All values are attainable, with the exception of the range of 1-9 points. Scores in the range of 1-9 points are censored at zero. A second reason why a project may receive a $pwert$ score equal zero is that that it received a "knock out" zero in one category.

During my interview with an examiner, it was emphasized to me that the evaluation is carried out at the most disaggragated level, i.e. at the level of technical characteristics. By requesting separate assessments and documentation at this level, the agency is deliberately trying to deter "point nudging", as a "holistic" assessment of the $pwert$ score is viewed to potentially give rise to the temptation to "push" borderline cases across the line.

The decision whether or not a funding application is approved rests with a board composed of political appointees. Based on all assessments, the examiners allot projects to the categories "recommended for funding", "not recommended for funding" and "to be discussed". There is no formal rule that stipulates how examiners arrive at a funding recommendation decision and there are no hard cut-offs. However, according to the interviewed examiner, projects with a $pwert$ score below 23 points are put into the category "not recommended for funding", unless they received very high assessments on the remaining dimensions. Projects with a $pwert$ score between 23 and 25 points are very often left "to be discussed" and projects with a higher $pwert$ score are usually "recommended for funding", unless they received very low assessments on the remaining dimensions. Projects in the category "to be discussed" are borderline cases. The board receives the expert assessments, descriptions of the projects and the funding recommendations, which are not binding for the board. In rare cases, the board overturns the funding recommendation (in either direction). The board makes funding decisions on the projects in the category "to be discussed". The exact rules under which the board operates and the transcripts of the board meetings are confidential.

Once an application is approved, the funding amount scales with no (absolute) upper limit to the total cost of the project. On average, applicants receive around 25% of the total cost of the project in cash ("free money").[6] The average funding amount was 153,000 Euro across all applications that were approved between 2002 and 2005. The funding does not come with special contractual provisions regarding patenting. Firms that receive funding are not required to patent the offspring of their research efforts and the creation of patents is not an explicit policy goal. Before I collected the data, the agency did not have patent records of the firms they were funding.

3.2 Data and construction of the baseline sample

The data from the funding agency FFG covers all applications to the "Basisprogramm" between 2002 and 2005. I manually match the set of FFG applicant firms to the set of patent applicants with an Austrian country code in the patent database PATSTAT and to the set of firms in the firm database AMADEUS, using the firm's name and address.

For each funding application, the $pwert$ score, the binary funding approval decision, the total cost of the submitted project, sector of the firm, firm age and year of application are available. I record the firm's patent applications, consolidated at the DOCDB patent family level. I distinguish between patent families that comprise an application at the European Patent Office (EPO), which I refer to as "EP Patents", and patent families that do no comprise an application at the EPO, which I refer to as "Non-EP Patents". In the Austrian context, the choice whether or not international coverage is sought reflects the perceived value of an invention.

[6]The matching rate of a project, defined as the funding amount divided by the total cost of the project, depends to a minor extent on firm and project characteristics. Figure 5 in the Appendix depicts the distribution of the matching rates for funded projects in the baseline sample. In Figure 6 in the Appendix, I provide plots of the matching rates for funded projects against the $pwert$ score. There is no evidence that matching rates for funded projects depend on the $pwert$ score.

The construction of the sample used for the analysis involves some exclusions from the full set of grant applications. In my research design, I relate patent applications by the firm after the funding application to the binary funding approval decision, using the $pwert$ score as an instrument. I exclude applications from firms with multiple funding applications in the same year. This is a limitation of this paper, since I exclude some of the most frequent recipients of FFG research grants. The remaining applications constitute 62.6% of the total number of applications, and the remaining applicant firms constitute 92% of the total number of applicant firms. As a robustness check, I consider the sample when I retain all firms and use the application that received the maximal $pwert$ across all applications of the firm in a given year.[7]

Descriptive statistics for the baseline sample, which comprises 2619 funding applications across 1936 firms, are presented in Table 1 (upper panel). I also present firm-level data from the agency on employment, R&D expenditures and sales. However, for more than one third of the applications in the baseline sample, this additional firm-level data is missing. It is missing in a systematic way: applications that were rejected and applications with low evaluation scores are more likely to have missing data. Therefore, I drop these variables from the baseline sample for my analysis. The distribution of the $pwert$ score, which appears symmetric around a peak at the median score of 30, is shown in Figure 1. In line with the characterization by the expert examiner, the peak is a feature of the underlying distribution of project quality and not evidence of manipulation. Category scores are placed on discrete grids of admissible integer values. Consequently, some $pwert$ scores are more frequent than others, because for some, there are more possible combinations of category scores that lead to same $pwert$ score than for others. Additional information for all variables is included in the Appendix.

4 Empirical strategy

My approach relies on the fact that funding approval is based on the main evaluation score $pwert$, and that there is a highly nonlinear relationship between the probability of funding approval and the $pwert$ score. The strategy leverages the intuition underlying a fuzzy regression discontinuity design and is very similar to the approach taken in Jacob and Lefgren (2011). Figure 2 shows the dependence of the probability of funding approval on the evaluation score $pwert$ for the baseline sample. The plot reveals a distinctive, nonlinear increase in the point range of 23 to 28. Given the institutional detail of the evaluation process, it appears reasonable to assume that the dependence of project quality on $pwert$ score does not exhibit a similar increase in this point range. Therefore, the distinct, steep increase of the funding approval probability in the score can be used to identify the causal effect of funding approval.

This empirical strategy is more likely to produce credible estimates of the causal effect of funding approval than the OLS estimate obtained from regressing patent outcomes on funding approval. The OLS estimate partly reflects the selection policy of the government agency and may overestimate or underestimate the

[7] Table 5 in the Appendix shows the number of applications per firm in each year between 2002 and 2005.

causal effect. If the agency systematically "rewards winners", the OLS estimate may overstate the effect of the research grant. On the other hand, if the agency is trying to "lend a helping hand" to firms that struggle or face particularly great challenges in research, the OLS estimate may have downward bias.

While the absence of an explicit "cutoff" makes the application of a standard fuzzy RDD infeasible, I may still leverage the distinct, nonlinear nature of the relationship between funding approval and $pwert$ score in an Instrumental Variables model. First, I model the steep increase of the funding approval probability in $pwert$ score with a nonlinear function $f_{\geq p*}(\tilde{\beta};\cdot)$, starting at point p^*. The location of break point p^* is chosen by best fit from data. I include linear spline functions on the ranges $[10, p^*]$ and $[p^*, 50]$ to control for linear increases of the funding probability in the evaluation score. In the main model, applications that received a $pwert$ score equal zero are excluded. The linear spline function starting at p^* is defined as $l_{\geq p^*}(pwert) = max(pwert - p^*, 0)$. The first stage equation is given by

$$approved_i = \alpha + \beta_1 pwert_i + \beta_2 l_{\geq p*}(pwert_i) + f_{\geq p*}(\tilde{\beta}; pwert_i) + \gamma X_i + \epsilon_i \tag{1}$$

where $approved_i$ is the binary funding approval decision and X_i is a vector of controls. Two models for the nonlinear increase $f_{\geq p*}(\tilde{\beta};\cdot)$ are considered: a model with discrete jumps and a model that uses a cubic spline.[8] Figure 2 displays the fit of the functions estimated in the first stage for both models. In the model that uses discrete jumps, the location of the discontinuities aligns with the description of the approval process by the examiner.

In the second stage, I exclude function $f_{\geq p*}(\tilde{\beta};\cdot)$ and control for the underlying relationship of project quality (directly affecting the outcome) and $pwert$ with a linear function on the entire score range and a linear spline starting at p^*. This specification allows for differences in the slope of the direct relationship of $pwert$ and outcome to the left and to the right of p^*. In an important robustness check, I omit the linear spline function starting at p^* and assume a constant slope in the relationship on the entire point range. The second stage is given by

$$EP\ Patent\ binary\ post_i = \alpha' + \tau\ \widehat{approved}_i + \beta_1' pwert_i + \beta_2' l_{\geq p*}(pwert_i) + \gamma' X_i + u_i \tag{2}$$

with orthogonality condition $f_{\geq p*}(\cdot; pwert_i) \perp u_i$. The outcome variable $EP\ Patent\ binary\ post$ is an indicator variable that measures whether or not the firm filed an application for a (for at least one) European Patent in the year of the funding application or during the subsequent three years after the funding application. The model relies on variation in a narrow point range of the $pwert$ score that produces big differences in funding approval probabilities to identify the effect. The model does not use the eventual funding approval decision for identification, which may reflect unobserved characteristics.

[8] In the model that uses three discrete jumps, $p^* = 23$ (selected by best fit) and $f_{\geq p*}(\tilde{\beta}; pwert_i) = \tilde{\beta}_3 \mathbb{1}_{\{pwert_i \geq 23\}}(pwert_i) + \tilde{\beta}_4 \mathbb{1}_{\{pwert_i \geq 26\}}(pwert_i) + \tilde{\beta}_5 \mathbb{1}_{\{pwert_i \geq 29\}}(pwert_i)$. The model with three jumps has considerably better fit than model that comprise one or two jumps. In the model that uses a cubic spline, $p^* = 21$ and $f_{\geq p*}(\tilde{\beta}; pwert_i) = \tilde{\beta}_3 l_{\geq 21}(pwert_i)^2 + \tilde{\beta}_4 l_{\geq 21}(pwert_i)^3$.

To help clarify this point, consider the following example: projects that receive a $pwert$ score of 25 are usually decided by the board, giving them a (roughly) 40 percent chance to be approved. For those projects, I do not assume that the ones that are picked by the board are comparable to those that are rejected. The board's funding decision may be based on characteristics that correlate with innovation, but which are not observed. In contrast, I assume, loosely speaking, that projects that received a score of 25 are comparable to projects that received a score of 26. Projects with this score are usually directly recommended for funding by the examiners, which gives them a considerably higher chance to be approved. The temptation for examiners to push their preferred projects "across the line" is counteracted by the institutional details of the evaluation process (discussed further below).

The exclusion restriction implies that it is assumed that there is no nonlinear increase in project quality coincidental with the increase in funding approval probability. In the context of the previous example, this means that, while there is a drastic increase in the chance to be approved moving from a score of 25 to a score of 26, the difference in project quality is assumed small. It is necessary to assume an inflexible functional form when I control for the direct dependence of the outcome variable on the $pwert$ score; otherwise, I would not able to seperate the effect of the nonlinear increase in funding approval from the effect of any potential nonlinear change in project quality.[9] For inference, I cluster standard errors at the firm level.[10]

Graphical evidence

The left column in Figure 3 shows how the binary outcome variable is fitted in the second stage regression for both models (without control variables). The curve in red depicts the direct dependence of the outcome variable on the $pwert$ score, given by $\hat{\alpha}' + \hat{\beta}_1' pwert + \hat{\beta}_2' l_{\geq p*}(pwert)$. The curve in black is the sum of the red curve and the predicted funding approval probability multiplied by the estimated treatment effect of funding, given by $\hat{\tau}\ \widehat{approved}(pwert) + \hat{\alpha}' + \hat{\beta}_1' pwert + \hat{\beta}_2' l_{\geq p*}(pwert)$. The larger the estimated treatment effect of funding, the larger the vertical distance between the black curve and the red curve.

Focusing on the black curve, I observe that, as I approach the point range of around 23 to 28 points from the right, the predicted propensity to file a European Patent declines. This suggests that projects of lower quality are less likely to lead to a patent application in this point range. In the point range of around 23 to 28 points, there is a sudden drop in the patenting propensity that coincides with the steep, nonlinear drop in the funding approval probability shown in Figure 2. The fact, that the patenting propensity declines at a much faster rate in this point range indicates that the loss of funding affects patenting.

[9]A limitation is that the estimated effect identified with this empirical strategy has a local interpretation. In Table 1 (lower panel), I present descriptives specifically for funding applications that fall in the point range of the $pwert$ score between 20 and 28 points. Table 1 (lower panel) also shows that such funding applications differ systematically from the rest of the sample.

[10]Kolesar and Rothe (2018) caution against clustering by the running variable in the context of regression discontinuity designs. When I cluster by $pwert$ score, I obtain standard errors that are always smaller than the standard errors that are reported in this paper.

Both models predict a significant, positive linear relationship of the patenting propensity and the $pwert$ score on the range of $[p^*, 50]$. In contrast, the relationship is insignificant on the range $[10, p^*]$. This could indicate that, if projects are rejected, they are not undertaken and a further reduction in project quality therefore does not further decrease the propensity to patent. In the right column of Figure 3, I present the specification where I assume a constant slope in the underlying direct relationship of outcome and score on the entire range $[10, 50]$ and omit the linear spline starting at p^*.

Specification of other control variables

The set of control variables comprises past European Patent and Non-EP Patent filings, project cost and fixed effects for firm age group, firm sector and year of application. Patent filings are included seperately as indicator functions and quadratic functions of the patent count. These variables capture “pre-sample” patent information and control for the firms’ fixed propensities to file patents, in the vein of Blundell et al. (1999). Project cost are included as a quadratic function. Definitions of all variables are included in the Appendix to Section 3.

Corroborating the exclusion restriction

There are two institutional features of the evaluation process that make it unlikely that the relationship between project quality and the $pwert$ score exhibits a similar distinctive, nonlinear increase as is visible in the dependence of funding approval on score (Figure 2). First, the fact that sub-scores are summed across separate categories makes it plausible that there are no discontinuous changes in the interpretation of the $pwert$ score in any particular range. Second, as described in section 3, precautions have been taken so that examiner do not sort “bad” projects into the point range just below the steep increase in funding approval, and “good” project just above it. The technical evaluation is carried out at the disaggregated level of technical characteristics, thereby further reducing the chance that examiners “nudge” applications into their preferred score range.

I conduct validity checks to strengthen confidence in the design. First, I implement a falsification test in which I apply the IV model to the 4 years preceding the funding application. A placebo effect in the period preceding the application would indicate that the effect cannot be fully attributed to the grant, but may reflect differences in time-invariant patent propensities of funded and unfunded firms. Figure 7 in the Appendix contrasts the estimated effects in the pre-grant period and in the post-grant period graphically (without control variables). The vertical distance between the black curve and the red curve is proportional to the estimated treatment effect. The regression results are shown in Table 6 in the Appendix. The estimates of the effect in the 4 years preceding the funding application appear small (range between -2.77 pp and 5.36 pp) and are insignificant.

Second, I try to detect sudden, nonlinear changes in the observable characteristics of projects and applicant firms across the point range in which the funding approval probability steeply increases. Table 7 and Table 8 in the Appendix present the results of applying the IV model to the control variables. Although I control for these variables in the main model, a significant difference identified by the model would cast doubt on its identifying assumption. In particular, if unobserved characteristics of projects were an important confounder, I would expect to find a significant difference in the project cost. I do not detect significant differences at the 0.05 level.[11]

The validity checks suggest that it is unlikely that unobserved confounding differences in firm characteristics or project characteristics explain the uncovered effects. In the next section, after presenting the baseline results, I corroborate the plausibility of the design by showing that differences in the patenting propensities of funded and unfunded firms are transient and eventually vanish.

5 Results

Effect of the grant on the propensity to apply for a European Patent

Table 2 presents the results for the effect of funding approval on the propensity to file a (at least one) European Patent in the year of the funding application or during the subsequent three years. I first discuss the estimates from the model that uses three discrete jumps as instrument (columns 1 and 2). The point estimate of the effect is 15.4 percentage points (pp) in the parsimonious specification without other control variables, which is statistically significant at the 0.05 level (column 1). When I include the full set of control variables, the estimated effect is revised by 0.43 standard deviations to 12.4 pp (column 2, statistically significant at the 0.05 level). I regard this estimate as the main result.

In the model that uses a cubic spline as instrument, the estimated effect of funding approval on the propensity to file a European Patent is 12.2 pp, which is significant at the 0.05 level (column 4). For reference, the mean of the dependent variable is 27.8 pp in the sample. Hence, my estimates suggest a substantial impact of grants on firms' patenting propensities. The results for all first stage regressions are shown in Table 9 in the Appendix.

Time delay of the effect and extended outcome period

I examine the delay of the treatment effect by studying the effect of funding approval on consecutive two-year periods after the funding application. Figure 4 shows point estimates and confidence intervals. I find the highest point estimate of 14.0 pp for years 3 and 4 after the funding application (year 0 being the year of the funding application). In contrast, the effect is only around 3.6 pp in years 1 and 2, and 4.9 pp in years 0 and 1. After year 5, the effect decays and is statistically indistinguishable from 0 ever after.

Hence, it appears that, in the majority of cases, firms need at least 3 years to develop their inventions

[11] I detect a significant difference at the 0.1 level for the firm age variable in the model that uses the cubic spline as instrument (Table 8 column 3). However, this finding appears to be driven by firms who delay incorporation in response to not receiving a grant (age can be negative for firms that incorporate after their funding application). The difference becomes insignificant when I winsorize the variable at the 95th percentile.

after receiving a grant. This finding alleviates concerns that the program simply induces firms to patent inventions that already existed, or to file patents in order to placate the agency. The grant is paid out within one year after funding approval, after which point the agency views the contract as fulfilled and is no longer in contact with the firm.

Following Bronzini and Piselli (2016) and Howell (2017), the main results are presented with patent outcomes measured for the year of the funding application and the subsequent three years (4 years total). In the Appendix, I show that the evidence becomes stronger when I extend the time frame over which outcomes are measured by one year (Table 10 columns 1 and 2, 15.4 pp and 13.9 pp, statistically signifcant at the 0.01 level). When I extend the outcome period further, the point estimates of the effect decline and become less precise (Table 10 columns 3-4, between 8.57 pp and 10.7 pp, which are marginally significant or insignificant).

The decaying, and eventually vanishing, effect after years 4 and 5 points to the transient nature of the grant. Funded firms are not set on different long-term trajectories, but revert in their patenting propensities to the level of unfunded firms. This finding appears inconsistent with a pattern where a one-time grant leads to a permanent increase in innovativeness via a change in organizational practices or via learning.

Robustness

First, I reestimate the model, assuming a constant slope in the direct linear relationship of the outcome and $pwert$ score on the entire range, illustrated in the right panel of Figure 3. Results are presented in Table 11 in the Appendix. Compared to the main result, the estimates are slightly revised downwards to 9.5 pp and 10.4 pp (columns 1 and 2, upper panel).

In the second robustness check, I maintain the specification of the main model, but include applications that received a $pwert$ score of zero. I obtain point estimates of 11.1 pp and 10.8 pp (Table 11 columns 3 and 4, upper panel). Across both robustness checks, the evidence becomes stronger when I extend the time frame over which outcomes are measured by one year (columns 1-4, lower panel).

Third, the construction of the baseline sample used in the analysis involves the exclusion of funding applications from firms that submitted multiple applications in the same year. When I instead retain all firms and use the application that received the maximal $pwert$ score across all funding applications of the firm in a given year, I find point estimates that range between 6.4 and 10.1 pp, shown in Table 12 in the Appendix, columns 1 and 2. The estimated effect may weaken because firms that submitted multiple applications, and which best application only marginally made it above the "quality-cutoff", had their other, weaker applications rejected.

Fourth, I constrain the bandwidth of the sample around the steep, nonlinear increase in the funding approval probability. Table 12 presents the results when I restrict the sample to applications in the range from

15 to 30 points or 20 to 30 points. The point estimates remain positive and range between 7.1 and 23.0 pp (columns 3-6, upper panel).[12] In an alternative specification, I try to leverage the steep, nonlinear increase in the funding approval probability in the most parsimonious way; I restrict the sample to applications that either fall in the point range of the $pwert$ score immediately before the increase in the funding approval probability or in the point range immediately after the increase. In this specification, I use an indicator variable as instrument and include no further controls for the $pwert$ score. The estimated effect of funding approval in this specification is 10.4 pp (Table 12 column 7, upper panel).

Last, I account for the uncertainty about the location of break points and jumps in a bootstrap in Table 13. I resample firms with replacement, choose break points and jumps for the obtained sample by best fit and then re-estimate the model. Except for the parsimonious model that uses three discrete jumps and does not include controls, the 95% confidence intervals of the estimates do not contain zero.

Additional Results: OLS estimates, Patent counts and Non-EP Patents

The estimates obtained using my empirical strategy contrast with the estimates obtained from an Ordinary Least Squares (OLS) regression of the propensity to file a European Patent on funding approval, presented in the Appendix in Table 14. The OLS estimate of the effect of funding approval in the model without controls is 17.9 percentage points (column 1), with a standard deviation of 1.87 percentage points. When I include the full set of control variables in the OLS model, the coefficient of funding approval drops by around 5.7 standard deviations to 7.3 percentage points (column 2). Columns 3 and 4 in Table 14 show the placebo effect in the 4 years preceding the funding application in the OLS model. Funding approval is highly correlated with the propensity to patent in the 4 years preceding the funding application. This suggests that the OLS estimates, in all likelihood, do not capture the causal effect. Persistent heterogeneity in firm characteristics, already present in the 4 years preceding the funding application, is a likely source of bias. The fact that the OLS estimate in column 2 is lower than the IV estimate indicates that, conditional on firm characteristics, the agency may select projects that are more likely to fail. This is consistent with the notion that the agency "lends a helping hand" for projects that pose a particularly great challenge.

Table 15 in the Appendix presents results for the effect of the grant on the number of EP patent applications in the year of the funding application or during the subsequent three years. I consider two models: a nonlinear IV model, which was suggested by Windmeijer and Santos Silva (1997), that assumes an exponential conditional mean in the outcome equation, and a linear IV model where I transform the outcome variable by the inverse hyperbolic sine. To account for the dispersed nature of patent counts, I winsorize patent count variables at the 99th or at the 95th percentile. Estimates of the (average marginal) effect in the nonlinear model range range between 0.49 and 0.97 additional EP patent applications per grant, which are either marginally significant or insignicant, depending on the model and the level of winsorization. In the

[12]However, in my setting the standard errors of the estimates become larger because, as the bandwidth narrows, it becomes more difficult to distinguish the linear change in project quality from the nonlinear increase in the funding approval probability. With a narrower bandwidth, the range of $pwert$ scores where the funding approval probability is "flat" becomes smaller relative to the range that exihibits the steep increase.

linear model, I find a marginally significant increase by around 14.1 to 16.3 percent (mean of dependent variable is 1.25 European Patents).

In Table 16 in the Appendix, I examine the effect on the propensity to apply for a Non-EP Patent in the year of the funding application or during the subsequent three years. To reiterate, in this study, Non-EP Patents are defined as patent families that do not comprise an application at the European Patent Office. This set of patent applications is typically filed at the Austrian Patent Office (only).[13] Non-EP Patents are common and account for 55.4 percent of patents in the baseline sample. However, the propensity to apply for a Non-EP Patent does not appear affected by the grant. The estimates of the effect range between -2.7 and 1.3 percentage points, and are not significant in any model. The mean of the dependent variable is 25.5 percent. I also examine the the effect on the number of Non-EP Patents and find estimates that are very close to zero and insignificant in all models (not shown). Because the decision whether or not international coverage is sought reflects the perceived value of the invention, this finding is consistent with the notion that grants increase the likelihood of high-value patents, but not the likelihood of low-value patents.

5.1 Heterogeneous treatment effects: experienced and young firms

Figure 8 and Figure 9 in the Appendix show the distribution of firm age in the baseline sample. I split the sample at the median age. I refer to firms that are older than 5 years as "experienced firms", and to firms that are at most 5 years old as "young firms". The results are presented in Table 3. For the model that uses three discrete jumps as instrument, the point estimate of the effect of funding approval is 21.0 percentage points for experienced firms, which is statistically significant at the 0.05 level. The point estimate of the effect for young firms is 4.21 percentage points, which is insignificant. Across all models, the treatment effect for young firms lies between 2.13 and 2.39 standard deviations below the treatment effect for experienced firms. The mean of the dependent variable is 33.4 percentage points for experienced firms and 23.5 percentage points for young firms. A back-of-the-envelope calculation suggests that young firms would have patented at a similar, but slightly higher rate than experienced firms in the absence any funding (20.8 percent compared to 18.9 percent). Placebo regressions on the period preceding the funding application for both groups of firms are presented in Table 17 in the Appendix. Robustness checks for alternative specifications of the model are included in Table 18 in the Appendix.

I further examine the heterogeneity of the effect for attributes that are potentially correlated with firm age. In Table 19 in the Appendix, I present split-sample estimates for the subsample of 2144 applications for which firm sales and number of employees of the firm are available. The lower panel of Table 20 presents results on heterogeneity based on the firm's R&D expenditures prior to the grant application, which are available for 2046 applications. I do not find evidence that larger or smaller firms, or firms with higher R&D expenditures, are particularly affected. In contrast, estimates of the effect for firms older than 5 years are large and statistically significant across all subsamples, while there is no discernible effect for firms younger than

[13]To be clear, a European Patent may first be filed with the Austrian Patent Office. However, it must additionally be applied for at the EPO within the priority year.

5 years. When I split the sample based on prior patenting experience, the estimated effect appears larger for firms that had filed at least one patent prior to the grant application (shown in the uppper panel of Table 20). In summary, the findings from this section suggest that grants to experienced firms have larger effects on the propensity to file a European Patent.

5.2 Do grants induce more ambitious research projects?

This section confronts claims, which are discussed in section 3.1, that grants induce projects that are technologically more ambitious than internally financed projects and allow firms to "push the envelope beyond [the firms'] normal type of R&D activity (Arnold et al. 2004, p. 54). The hypothesized effects of grants are supposedly pronounced for large and experienced firms (Arnold et al. 2004, p. 52).

I operationalize the suggested effects alluded to by Arnold et al. 2004 with two backwards-citation measures that gauge the knowledge that patents draw on. First, I employ an unconventionality measure proposed by Berkes & Gaetani (2020), that quantifies how atypical combinations of cited technological classes are when they appear together on a patent. Berkes & Gaetani (2020) show that unconventional patents have a disproportionate chance of becoming widely cited "hit" patents. Second, I determine whether a patent cites a technological class that the firm had never cited prior to the funding application.

First, I assign patents their percentile in the unconventionality score distribution relative to a comparable population of patents in PATSTAT, based on their cited 2-digit IPC classes. 56.2% of patents in the baseline sample (including patents prior to and after the funding application) reside in the less conventional, i.e more unconventional, half of the unconventionality score distribution. 9.1% are in the top decile. In this section, I do not differentiate between patents filed with the European Patent Office, and patents filed with National Offices. In Table 4, I apply the IV model that uses three discrete jumps to the propensity to file a patent in the bottom half of the unconventionality distribution (column 1 and 2), which I refer to as a "conventional patent", and to the propensity to file a patent in the upper half of the unconventionality score distribution (columns 3 and 4). I consider the full sample of all applicant firms, and the sample of experienced firms above the median firm age of 5 years, where I expect effects to concentrate.

Estimates of the effect of the grant on the propensity to file a conventional patent are small (3.79 pp and 3.37 pp) and insignificant for all applicant firms and for experienced firms. In contrast, estimated effects for patents that fall in the upper half of the unconventionality score distribution are considerably larger (11.1 pp and 19.2 pp), and are either marginally significant or significant at the 0.05 level. In columns 5 and 6, I examine patents that fall in the top decile of the distribution. The grant appears to increase the propensity to file a highly unconventional patent for experienced firms by 10.3 pp, which is marginally significant. In the sample of all applicant firms, the effect of the grant on the propensity to file a highly unconventional patent is only 2.0 pp, which is insignificant.

Second, I compare the IPC classes cited by a patent to the IPC classes that are cited in earlier patents of the firm. When I exclude applications by firms without previously cited IPC classes (prior to the funding application), I retain 971 applications across 685 firms. In this sample, 28.6% of patents contain citations of technology classes that are "novel to the firm". Estimates of the effect on the propensity to file a patent that does not have this property, shown in columns 7 and 8, are close to zero (4.1 pp for all firms and -1.4 pp for experienced firms). In contrast, estimates of the effect on the propensity to file a patent that cites a technological class that is novel to the firm appear large (17.3 pp for all firms and 20.6 pp for experienced firms, columns 9 and 10). However, these estimates are marginally insignificant at the 0.1 level (p=0.16 for all firms and p=0.13 for experienced firms), presumably due to the loss of power when restricting the sample. In the Appendix, I show that there is a statistically significant correlation between the funding approval decision and the "average number of novel technological classes" per patent filed in the outcome period.

6 The Model

This section proposes a model that reconciles the institutional features of the program and the evidence. In this model, grants affect the behavior of firms by altering the choice between competing research projects, in the vein of Buisseret, Cameron and Georghiou (1995). Internal competition from more conventional projects within firms explains why some more ambitious and explorative projects may not be undertaken in the absence of a grant, even and especially in experienced or potentially well-resourced firms. The structure of the model echoes the supposition in Buisseret, Cameron and Georghiou (1995), that "[t]ypically the company draws up a list of projects [..] and reviews the available sources of public funding to see whether it can match them to the list. Additionality in this context is best seen the at portfolio [..] level." (p. 598).[14] The firm has the option to undertake an ambitious project that explores a line of research that is new to the firm, which is inherently risky. If successful, it spawns a series of follow-up projects that can be exploited subsequently. Hence, firms incessantly choose between exploiting their existent research lines and exploring new research lines.

The policy of the agency is rationalized in this model by an externality that is largest for the initial breakthrough. The non-appropriable social value of subsequent incremental projects is assumed to decline gradually. A possible justification for this assumption is that explorative research is more likely to produce knowledge gains that are diffuse or distant from the firm's core competence, and hence may be difficult to exploit commercially. Furthermore, conditional on the firm already having implemented several projects that follow the same line of research, the subsequent incremental improvements are likely specialized and not applicable by others, as they typically lack relevant prior knowledge. This position is indirectly supported by evidence presented by Berkes and Gaetani (2020), who show that patents with unconventional knowledge inputs are disproportionately likely to become "hit" patents that instigate follow-on research, as evidenced by forward patent citations.

[14] Buisseret, Cameron and Georghiou (1995) further claim that, in their portfolio model, it can be rationalized that public funding induces a shift towards riskier projects "further down the list of priorities" and that public funding resolves the problem that firms avoid long-term projects in favor of short-term projects (p. 593).

The firm decision problem

There is one firm in this dynamic decision problem. Initially, the firm has no existent research line. In each period, the firm receives an idea for a novel research line (with probability $\tilde{p} = 1$). Exploration costs $c_e > 0$ and yields a "success" with probability $0 < p_e < 1$. We may think that with probability $1 - p_e$, the firm either may make a discovery that it cannot commercialize, or that the project may simply "fail". If the firm decides to explore and succeeds in establishing a research line, it earns a stage-payoff of K and becomes a firm with a research line of vintage 1 in the next period. All payoffs are discounted at the rate $0 < \beta < 1$ and there is no terminal period. If I denote the value of a firm with no existent research line by V^o and the value of a firm with a research line of vintage 1 as V_1, the Bellman equation of the firm is

$$V^o = max\Big\{\beta V^o; -c_e + p_e(\beta V_1 + K) + (1 - p_e)\beta V^o\Big\} . \tag{3}$$

A newly established research line spawns a finite number of incremental projects N that can be researched in subsequent periods. I call the number of projects that have already been worked on the same research line the "vintage of the research line". The next incremental project of a research line of vintage T delivers payoff $K\lambda^T$, where $\lambda < 1$, succeeds with very high probability $p_i > p_e$ and costs c_i. I assume that the subsequent research projects are technically closely related and therefore entail a lower risk of failure. To simplifiy the analysis, I assume that $p_i = 1$. I assume that all incremental projects have a positive net present value so that even for the last project $K\lambda^N > c_i$. Furthermore, I require that $K + \sum_{t=1}^{N} \beta^t(K\lambda^t - c_i) > c_e/p_e$. This implies that the total discounted value of a research line is sufficient to motivate some research.[15]

During the subsequent periods, the firm could, in principle, work through the entire series of incremental projects on the existent research line. However, the firm continues to receive ideas for yet other research lines (with probability $\tilde{p} = 1$). As before, exploring a novel research line costs c_e and succeeds with probability p_e. If successful, the new research line fully replaces the old research line. The old research line is assumed to be a fall back. A firm with a research line of vintage T compares the value of exploring a novel research line with the continuation value of exploiting the currently active research line of vintage T. Consider, for example, a car producer trying to reduce the gas mileage for the next product cycle of a car model: the current technology of the engine can be improved incrementally to reduce gas mileage from 3 g/hm to 2.8 g/hm for the next generation of the model. The car producer is aware of a novel technology that could reduce gas mileage to 2.5 g/hm, but researching this technology is risky. If the car producer decides to explore the novel technology and succeeds, there is no point in undertaking the incremental project on the old technology and all cars in the next product cycle will be equipped with the novel technology. If the car producer decides to explore the novel technology and fails, she can still implement the incremental improvement on the current technology. The Bellman equation is

$$V_T = max\Big\{K\lambda^T - c_i + \beta V_{T+1}; -c_e + p_e(K + \beta V_1) + (1 - p_e)(K\lambda^T - c_i + \beta V_{T+1})\Big\} . \tag{4}$$

[15] Under this assumption, I show in the proof of Proposition 1 that undertaking the project must be optimal for firms that have no active research line.

If the firm explores and succeeds in establishing the new research line, it earns stage-payoff K, the old research line is retired and the vintage of the currently active research line is reset. Equation (4) gives rise to the condition that exploration is optimal for an firm with a research line of vintage T if and only if

$$\frac{c_e}{p_e} - c_i \leq \beta(V_1 - V_{T+1}) + K - K\lambda^T \; . \tag{5}$$

Exploration is optimal if the gains from renewing the research line outweigh the cost, taking into account the risk of failure. I assume that if a firm exhausts all incremental projects on its currently active research line, it becomes a firm with no existent research line and its value is given by V^o.

The following propositition describes the behavior of the firm in the firm decision problem. When a firm introduces a new research line, it exploits the research line for a few periods and eventually switches to exploring another research line.

Proposition 1. ***(Firm Behavior)*** *Consider the firm decision problem defined above. There exists a threshold $\bar{T}$ such that a firm explores a novel research line if the vintage of its existent research line T exceeds $\bar{T}$. Otherwise, it undertakes the incremental project.*

Proof. See the Appendix. □

The Social Planner Problem

I assume that there is a scaling factor $S > 1$ so that the total social payoffs of projects are SK, $SK\lambda$, $SK\lambda^2$ et cetera. The non-appropriable social value of projects is consequently given by $(S-1)K, (S-1)K\lambda, (S-1)K\lambda^2$. This tractable payoff structure has the property that the value not appropriated by the firm declines as research lines become more dated, and is largest for the explorative project that establishes a new research line.

I define the social planner problem as the firm decision problem with all project payoffs scaled by factor S, while costs and discount factor are left unchanged. The finding from Proposition 1 is still valid in the social planner problem. I introduce the assumption that the cost of the explorative research project c_e and the cost of the incremental research project c_i satisfy $\frac{c_e}{p_e} > c_i$.

This setting rationalizes the use of matching grants that target risky projects as the most cost-effective way to induce exploration.

Proposition 2. ***(Social Planner Solution and First-Best Grant Policy)*** *Consider the firm decision problem defined in the previous section and the social planner problem defined above. Let $\bar{T}$ be the threshold established in Proposition 1, such that the firm explores if the vintage of its existent research line exceeds $\bar{T}$. Let $\bar{T}^*$ be the corresponding threshold in the social planner problem. Then,*

- *the social planner explores new research lines earlier, i.e.* $\bar{T}^* \leq \bar{T}$

Furthermore,

- *the socially optimal research policy is implemented by offering targeted matching grants to firms with research lines of vintage* T *s.t.* $\bar{T}^* \leq T < \bar{T}$ *for explorative projects with matching rates* τ_T^{RG} *that solve* $\frac{c_e(1-\tau_T^{RG})}{p_e} - c_i = \beta(V_1 - V_{T+1}) + K - K\lambda^T$

where all value functions are obtained from the firm decision problem. The policy minimizes the expected sum of discounted subsidy payments to firms across all policies that implement the socially optimal research policy and that leave the firm at least as good off as in the firm decision problem (participation constraint).

Proof. See the Appendix. □

The policy of Proposition 2 abstracts from compliance issues, informational limitations or legal constraints that real world agencies may face. Under this policy, all grants are effective, in the sense that they induce a change in the firm's behavior relative to the status-quo of not receiving a grant.

Grant effect with fixed matching rate and imperfect compliance

I consider a modification of the firm decision problem that resembles the actual policy of the agency. In contrast to the previous section, I take the policy of the agency as given. In each period, with exogenous probability $0 < p_g < 1$, the firm receives a grant that lowers the cost of the explorative project by the fixed matching rate $\tilde{\tau}^{RG}$ to $c_e(1 - \tilde{\tau}^{RG})$. With probability $(1 - p_g)$, it does not receive a grant.

I assume that, if the firm receives a grant, it may divert the funding amount. For example, the firm may use grant money for the incremental project and misrepresent the expense to the agency. In this case, the firm derives a private benefit of $Dc_e\tilde{\tau}^{RG}$. I assume that diversion is costly or involves a reputational risk with the agency, so that $D < 1$. If the firm receives a grant, the condition that determines whether a firm with a research line of vintage T undertakes the explorative project is given by

$$\frac{c_e\big(1 - (1 - D)\tilde{\tau}^{RG}\big)}{p_e} - c_i \leq \beta(\tilde{V}_1 - \tilde{V}_{T+1}) + K - K\lambda^T , \tag{6}$$

where $\tilde{V}$ is the firm's value function prior to learning whether or not it receives a grant. If the firm does not receive a grant, the condition for undertaking the explorative project is

$$\frac{c_e}{p_e} - c_i \leq \beta(\tilde{V}_1 - \tilde{V}_{T+1}) + K - K\lambda^T . \tag{7}$$

There are two potential reasons why a grant might not induce a change in behavior: first, the firm may have undertaken the explorative project even in the absence of a grant. This corresponds to the situation where both condition (6) and condition (7) are satisfied. Second, the firm may not undertake the explorative project, corresponding to the situation where condition (6) is not satisfied. I refer to a grant as "effective" if the firm undertakes the explorative project when it receives a grant, but does not do so if it is denied funding. In congruence with the empirical part, causal effects are identified by comparing behavior in the event of receiving a grant with behavior in the counterfactual event of not receiving a grant.

Proposition 3. ***(Grant Effect and Firm Behavior)*** *Consider the modification of the firm decision problem defined above, where firms receive a grant with probability p_g that lowers the cost of the explorative project by the fixed matching rate $\tilde{\tau}^{RG}$ and firms may divert funding. Then, there exist thresholds $\tilde{T}_{RG} \leq \tilde{T}_{\neg RG}$ such that,*

- *for firms with an existent research line of vintage T:*
  - *if $T < \tilde{T}_{RG}$: the grant is ineffective; funded firms undertake the incremental project*
  - *if $\tilde{T}_{RG} \leq T < \tilde{T}_{\neg RG}$: the grant is effective; funded firms undertake the explorative project*
  - *if $\tilde{T}_{\neg RG} \leq T$: the grant is ineffective; funded firms undertake the explorative project, but would have done so even if they had not received a grant*
- *for firms with no existent research line, the grant is ineffective; funded firms undertake the explorative project, but would have done so even if they had not received a grant*

Proof. See the Appendix. □

In principle, the socially optimal research policy described in Proposition 2 can be implemented if $\tilde{\tau}^{RG}$ is selected such that $\tilde{T}_{RG} = \bar{T}^{*}$. However, there is waste.

I relate the age of firms, measured by the years since their incorporation, to the presumed vintage of their research lines. I assume that the vintage of a research line corresponds to the years since it was established. The transition probabilities between different vintages across years, as implied by Proposition 3, can be represented with the Markov transition matrix **M**, which is described in the Appendix. I denote the firm distribution over different vintages of research lines $\{o, 1, 2, 3, ..., N\}$ by a row vector f and assume that firms of age cohort 0 are all initialized in state o without an active research line, so that $f_0 = (1, 0, 0, .., 0)$. The distribution over vintages of research lines for firms of age cohort S is then given by $f_S = f_0 \cdot \mathbf{M}^S$. The share of firms in age cohort S for which the grant is effective is given by $\pi_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = f_S \cdot d(\tilde{T}_{RG}, \tilde{T}_{\neg RG})$, where d is a column vector that takes value 1 in all vintages T s.t. $\tilde{T}_{RG} \leq T < \bar{T}_{\neg RG}$ and 0 otherwise. I define the average causal effect of the grant on patenting outcome y for firms in age cohort S as $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = \pi_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG})\Delta y$. Δy denotes the difference in the patenting outcome when the explorative project is undertaken instead of the incremental project. I assume that $\Delta y > 0$.

Proposition 4. ***(Grant Effect by Firm Age)*** *Consider the modification of the firm decision problem. Let $\tilde{T}_{RG}$ and $\tilde{T}_{\neg RG}$ be the thresholds derived in Proposition 3 and suppose that $\tilde{T}_{RG} \neq \tilde{T}_{\neg RG}$. Let $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG})$ be the average causal effect of the grant on the patenting outcome for firms in age cohort S. Then,*

- *for firms in age cohort $S < \tilde{T}_{RG}$, $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = 0$.*

- *for firms in age cohort $S \geq \tilde{T}_{RG}$, there is an effect size $\tau^* > 0$ and a threshold firm age $\bar{S} \geq \tilde{T}_{RG}$ such that $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) \geq \tau^*$ for all $S \geq \bar{S}$.*

Proof. See the Appendix. □

Proposition 4 predicts that for experienced firms, grants trigger a change in behavior and affect patenting outcomes. The absence of an effect for young firms is based on a combination of two mechanisms: first, because the research line of a young firms is not dated, its continuation value has not yet declined to the point where the incentive provided by the grant eclipses the value of exploitation. Second, young firms are more likely to have not yet established a research line, which implies that they are incentived to undertake the explorative project irrespective of whether or not they receive a grant.[16]The average causal effect $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG})$ is 0 for the most recent firm cohorts and converges to a strictly positive number as cohorts age. However, a limitation of Proposition 4 is that convergence is not necessarily monotonic. As a consequence, comparisons of the grant effect for firms below and above a set age depend on the shape of the underlying firm age distribution.

7 Conclusion

This paper presents evidence from the Austrian Research Promotion Agency FFG of a substantial effect of government research grants on firm patenting outcomes. The grants entail considerable co-financing, as they fund only 25% of project cost, but scale to the size of projects without placing an upper limit on cost. My estimates suggest that a government research grant increases the propensity to file a patent application with the European Patent Office within 4 years by around 12 percentage points. The average grant size awarded by the agency is 118,000 Euro. The effect is particularly strong for firms above the median age, which is 5 years. On the other hand, estimates of the effect for younger firms are economically small and statistically insignificant. I present evidence that research grants encourage experienced firms to file unconventional patents and to cite technology classes that they had not cited previously.

Consistent with this evidence, I propose a model in which internal competition from conventional projects within firms explains why some more ambitious projects may not be undertaken in the absence of a grant, especially in experienced firms. By offering targeted research grants for risky projects, the agency encourages firms to undertake explorative research that may lead to novel research lines, as opposed to incremental research on old research lines. The model predicts that, as firm-cohorts age, there is a strictly positive share of older firms whose behavior changes due to the grant.

The findings of this paper deviate from other studies that find a weaker or insignificant effect for grants paid to older firms (Howell 2017, Santoleri et al. 2021). The grants examined in these papers, which are of fixed size and do not require co-financing, appear to alleviate financial frictions. In contrast, this paper points to a different mechanism for the research grants paid by the Austrian Research Promotion Agency.

[16]In reality, this may be an instance in which grants affect outcomes by alleviating financial frictions.

The R&D support programs of other EEA-member states, such as Tekes (Finland), Vinnova (Sweden), IWT (Belgium) or RCN (Norway), operate in a context that appears the most comparable. The results and theoretical considerations presented in this paper are potentially relevant for them.

The design of research grants may shape which barriers to innovation are most effectively addressed. Inefficiency in the direction of research may be especially relevant for experienced firms. Over the last decades, directed subsidies have been replaced by undirected R&D tax credits in many OECD countries (see OECD 2014). However, due to their indiscriminate nature, tax credits inadequatly affect the selection of research projects within firms. Such considerations, among others, suggest that government research grants are potentially important tools of innovation policy. Further research to tease out empirical findings across programs, and related theoretical considerations, are important avenues for further work.

References

[1] Arrow, K., 1962. Economic Welfare and the Allocation of Resources for Innovations. R. Nelson (ed.), The Rate and Direction of Inventive Activity, Princeton University Press.

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

[3] Arnold, E., Falk, M., Falk, R., Knoll, N., Leo, H., & Schwarz, G., 2004. Evaluation of the Austrian Industrial Research Promotion Fund and the Austrian Science Fund.
URL: www.ffg.at/sites/default/files/downloads/page/bpevaluierung03synthesisrep.pdf

[4] Becker, B., 2015. Public R&D policies and private R&D investment: a survey of the empirical evidence. Journal of Economic Surveys, 29(5), 917-942.

[5] Bérubé, C. & Mohnen, P., 2009. Are firms that receive R&D subsidies more innovative? Canadian Journal of Economics, 42(1), 206-225.

[6] Berkes, E. & Gaetani, R., 2021. The Geography of Unconventional Innovation. The Economic Journal, 131(636), 1466–1514.

[7] Besanko, D., Tong, J. & Wu, J., 2018. Subsidizing research programs with "if" and "when" uncertainty in the face of severe informational constraints. RAND Journal of Economics, 49(2), 285-310.

[8] Bloom, N., Schankerman, M., & Van Reenen, J., 2013. Identifying Technology Spillovers and Product Market Rivalry. Econometrica, 81(4), 1347-1393.

[9] Bloom, N., Griffith, R., & Van Reenen, J., 2002. Do R&D tax credits work? Evidence from a panel of countries 1979–1997. Journal of Public Economics, 85(1), 1-31.

[10] Blundell, R. & Griffith, R., & Van Reenen, J., 1999. Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms. Review of Economic Studies, 66(3), 529-554.

[11] Bronzini, R. & Iachini, E., 2014. Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach. American Economic Journal: Economic Policy, 6(4), 100-134.

[12] Bronzini, R. & Piselli, P., 2016. The impact of R&D subsidies on firm innovation. Research Policy, 45(2), 442-457.

[13] Bryan, K. & Lemus, J., 2017. The direction of innovation. Journal of Economic Theory, 172, 247-272.

[14] Buisseret, T.J., Cameron, H.M. & Georghiou, L., 1995. What difference does it make? Additionality in the public support of R&D in large firms. International Journal of Technology Management, 10, 587-600.

[15] Cappelen, A., Raknerud, A. & Rybalka, M., 2012. The effect of R&D tax credits on patenting and innovations. Research Policy, 41(2), 334-345.

[16] Clarysse, B., Wright, M. & Mustar, P., 2009. Behavioural additionality or R&D subsidies: A learning perspective. Research Policy, 38(10), 1517-1533.

[17] Czarnitzki, D., Hanel P. & Rosa, JM, 2012. Evaluating the impact of R&D tax credits on innovation: A microeconometric study on Canadian firms. Research Policy, 40(2), 217-229.

[18] Czarnitzki, D. & Lopes-Bento, C., 2013. Value for money? New microeconometric evidence on public R&D grants in Flanders. Research Policy, 42(1), 76-89.

[19] Dechezleprêtre, A., Einiö, E., Martin, R., Nguyen, K. T. & Van Reenen, J., 2023. Do tax incentives for research increase firm innovation? An RD design for R&D. American Economic Journal: Economic Policy (forthcoming).

[20] Falk, R., 2007. Measuring the effects of public support schemes on firms' innovation activities. Survey evidence from Austria. Research Policy, 36(5), 665-679.

[21] Demeulemeester, S., Hottenrott, H., 2017. R&D Subsidies and Firms' Cost of Debt. Working paper.

[22] Einiö, E., 2014. R&D subsidies and company performance: Evidence from geographic variation in government funding based on the ERDF population-density rule. Review of Economics and Statistics, 96(4), 710-728.

[23] Georghiou, L., 2002. Additionality and impact of R&D subsidies. IWT Studies, 40 (2002), pp. 57-64.

[24] González, X., Jaumandreu, J. & Pazó, C., 2005. Barriers to innovation and subsidy effectiveness. RAND Journal of Economics, 36(4), 930-950.

[25] Hall, B. H., 2002. The financing of research and development. Oxford Review of Economic Policy, 18(1), 35-51.

[26] Hall, B. H., Jaffe, A. B. & Trajtenberg, M., 2005. Market Value and Patent Citations. RAND Journal of Economics, 36(1), 16-38.

[27] Howell, S., 2017. Financing Innovation: Evidence from R&D Grants. American Economic Review, 107(4), 1136–1164.

[28] Hünermund, P. & Czarnitzki, D., 2019. Estimating the causal effect of R&D subsidies in a pan-European program. Research Policy, 48(1), 115-124.

[29] 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.

[30] Jacob, B. A. & Lefgren, L., 2011. The impact of research grant funding on scientific productivity. Journal of Public Economics, 95, 1168-1177.

[31] Jaffe, A. B., 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value. American Economic Review, 76(5), 984-1001.

[32] Jaffe, A. B., 1998. The Importance of "Spillovers" in the Policy Mission of the Advanced Technology Program. Journal of Technology Transfer, 23(2), 11-19.

[33] Jones, C. & William, J., 1998. Measuring the Social Rate of Return to R&D. Quarterly Journal of Economics, 113(4), 1119-1135.

[34] Lach, S., 2002. Do R&D subsidies stimulate or displace private R&D? Evidence from Israel. Journal of Industrial Economics, 50(4), 369-390.

[35] Lach, S., Neeman, Z. & Schankerman, M., 2017. Government Financing of R&D: A Mechanism Design Approach. American Economic Journal: Microeconomics, 13(3), 238-272.

[36] Lerner, J., 2000. The government as venture capitalist: the long-run impact of the SBIR program. The Journal of Private Equity, 3(2), 55-78.

[37] Lerner, J., 2009. Boulevard of broken dreams: why public efforts to boost entrepreneurship and venture capital have failed-and what to do about it. Princeton University Press.

[38] Lichtenberg, F. R., 1988. The private R&D investment response to federal design and technical competitions. American Economic Review, 78(1), 550-559.

[39] Mamuneas, T. P., & Nadiri, M. I., 1996. Public R&D policies and cost behavior of the US manufacturing industries. Journal of Public Economics, 63(1), 57-81.

[40] Neicu, D., Teirlinck, P. & Kelchtermans, S., 2015. Dipping in the policy mix: Do R&D subsidies foster behavioral additionality effects of R&D tax credits? Economics of Innovation and New Technology, 25(3).

[41] Nelson, R. R., 1959. The Simple Economics of Basic Scientific Research. Journal of Political Economy 67(3), 297-306.

[42] OECD, 2014. Science, Technology and Industry Outlook 2014.

[43] Rao, N., 2016. Do tax credits stimulate R&D spending? The effect of the R&D tax credit in its first decade. Journal of Public Economics, 140, 1-12.

[44] Santoleri, P., Mina, A., Di Minin, A. & Martelli, I., 2022. The causal effects of R&D grants: evidence from a regression discontinuity. The Review of Economics and Statistics, 1-42.

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

[46] Takalo, T., Tanayama, T. & Toivanen O., 2013b. Market failures and the additionality effects of public support to private R&D: Theory and empirical implications. International Journal of Industrial Organization, 31(5), 634-642.

[47] Trajtenberg, M., 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. RAND Journal of Economics, 21(1), 172-187.

[48] Wallsten, S.J., 2000. The effects of government-industry R&D programs on private R&D: the case of the small business innovation research program. RAND Journal of Economics 31(1), 82-100.

[49] Windmeijer, F.A.G., Santos Silva, J.M.C., 1997. Endogeneity in Count Data Models: An Application to Demand for Health Care. Journal of Applied Econometrics, 12, 281-294.

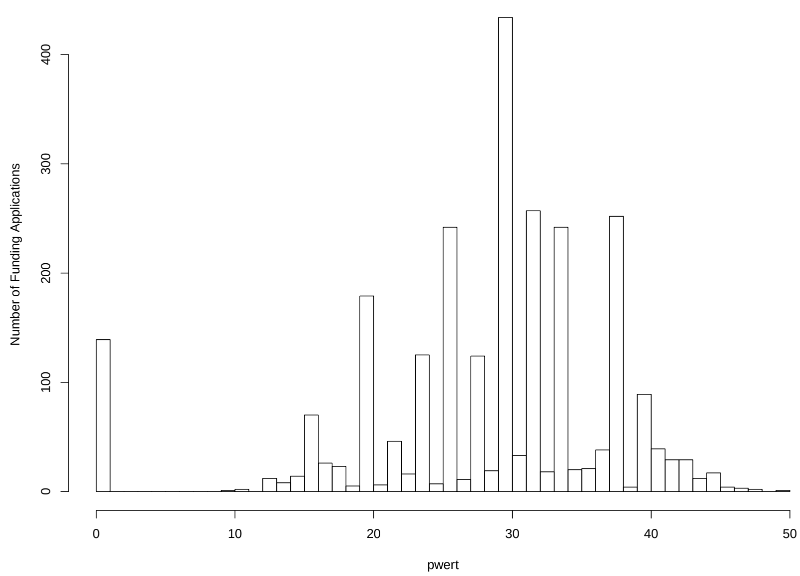
[50] Zúñiga-Vicente, J., Alonso-Borrego, C., Forcadell, F.J., Galàn, J.I., 2014. Assessing the effect of public subsidies on firm R&D investment: a survey. Journal of Economic Surveys 28(1), 36-67.

Table 1: Descriptive statistics for applications and applicant firms

	Project cost in Thsd. Euro	R&D Exp in Thsd. Euro	#Employees	Sales in Thsd. Euro	Firm Age	#EP Patents 12 yrs prior	#EP Patents 4 yrs after	#Non-EP Patents 12 yrs prior	#Non-EP Patents 4 yrs after
Baseline sample of applications									
# data available	2,619	2,046	2,144	2,144	2,444	2,619	2,619	2,619	2,619
Mean	529.21	1241.37	119.27	29682.88	11.68	1.85	1.22	2.68	1.03
Std. Dev.	1036.88	7127.42	454.39	206359.50	18.20	7.57	4.28	23.61	4.51
.10 Percentile	31.50	5.00	2	84.00	-2	0	0	0	0
.25 Percentile	130.00	50.00	6	490.00	1	0	0	0	0
.50 Percentile	276.94	213.00	21	2797.50	6	0	0	0	0
.75 Percentile	535.60	741.00	97	16477.50	16	1	1	1	0
.90 Percentile	1062.26	2487.00	310	61074.00	31	4	3	5	3
Applications around approval threshold (point range 20 to 28)									
# data available	756	576	607	607	702	756	756	756	756
Mean	333.74	560.37	89.18	14210.98	10.59	0.73	0.48	1.10	0.49
Std. Dev.	453.00	2689.41	700.06	49374.03	14.57	2.52	2.06	3.56	1.56
.50 Percentile	217.03	129.00	15	1570.00	6	0	0	0	0
Applications further from approval threshold (not in point range 20 to 28)									
# data available	1,863	1,470	1,537	1,537	1,742	1,863	1,863	1,863	1,863
Mean	608.54	1509.60	131.16	35793.13	12.12	2.31	1.52	3.33	1.26
Std. Dev.	1186.01	8223.99	306.94	241535.7	19.46	8.79	4.87	27.88	5.24
.50 Percentile	304.10	290.00	27	3400	6	0	0	0	0
t-statistic Difference in means	-6.19	-2.71	-1.93	-2.18	-1.88	-4.87	-5.64	-2.19	-3.96

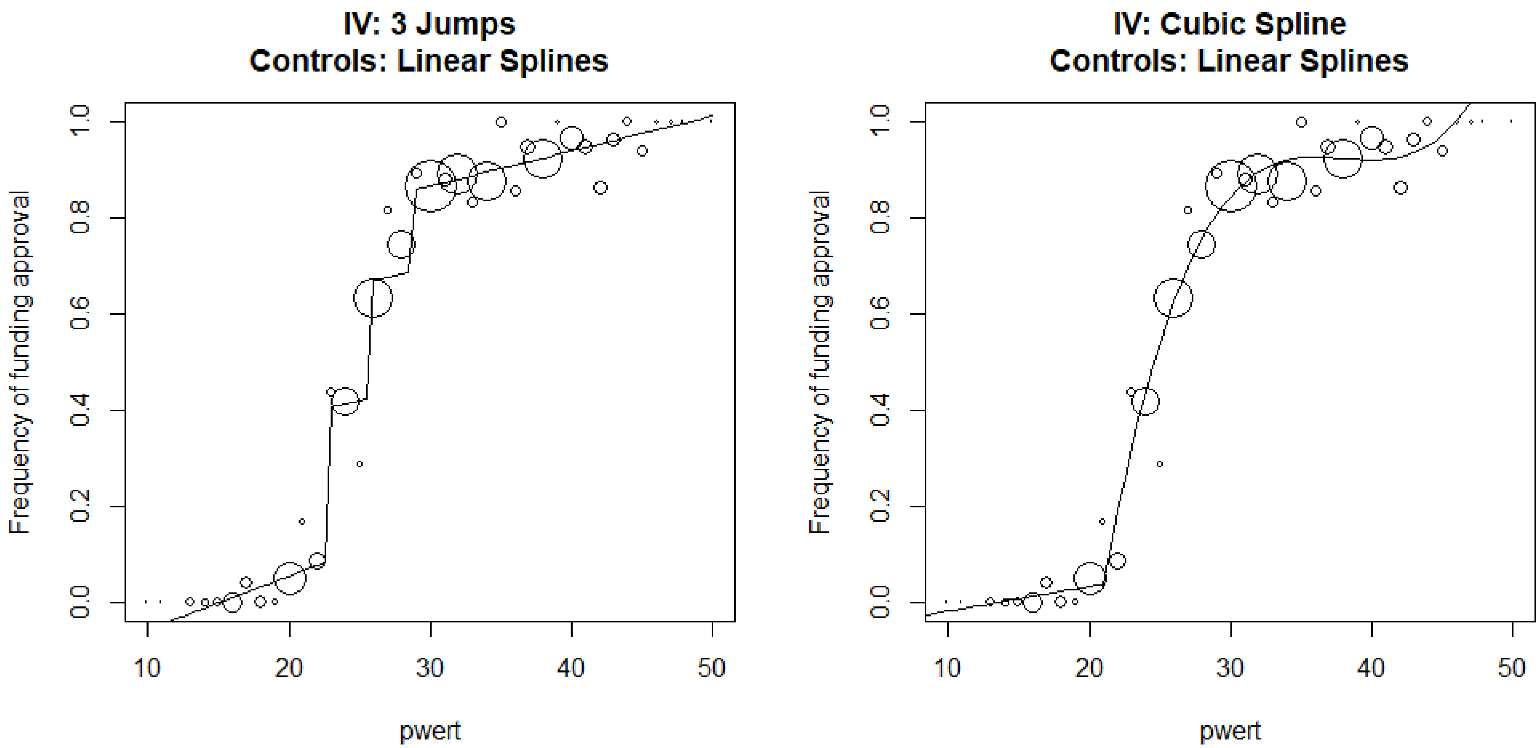
Note: Table presents statistics for the 2,619 funding applications (across 1,936 distinct firms) which are included in the baseline sample. 66 percent of funding applications were approved by the agency. The distribution of matching rates (funding amount divided by project cost), conditional on funding approval, is shown in Figure 5 in the Appendix. Share of applicant firms with at least one European Patent (Non-EP Patent) in the 12 years prior to the funding application is 0.299 (0.303). Share of applicant firms with at least one European Patent (Non-EP Patent) in the 4 years after the funding application is 0.278 (0.249). R&D expenditures, number of employees and sales are measured in the year prior to the year of the funding application, whenever available. A negative firm age indicates that the firm incorporated after the funding application.

Figure 1: Distribution $pwert$ score in the baseline sample



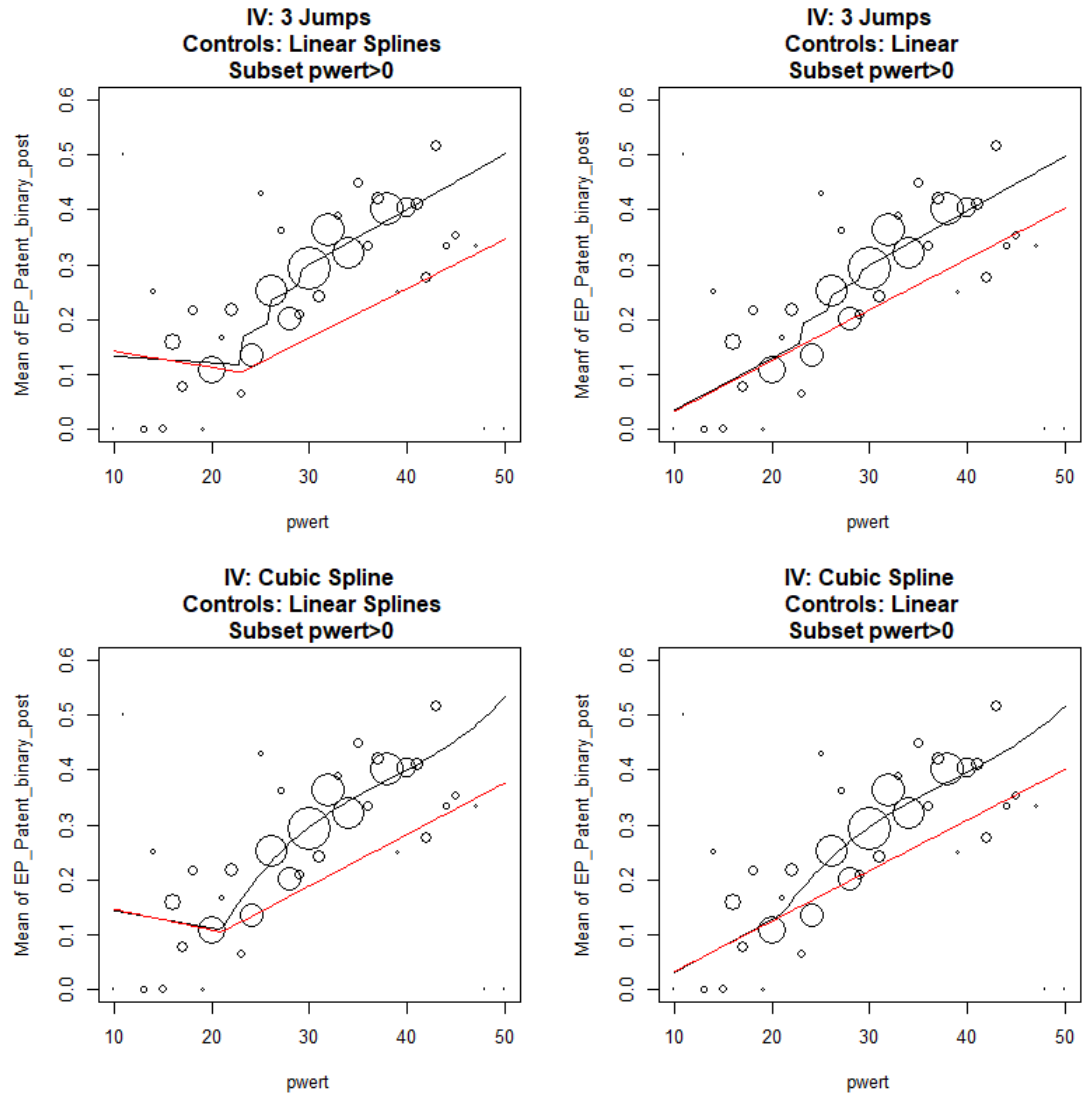
Note: The $pwert$ score is the evaluation score for the technical quality of the project, which is described in section 3. Dependence of funding approval on $pwert$ score is shown in Figure 2.

Figure 2: Fit for First Stage Regression: Funding approval probability



Note: Graphs show the frequency of funding approval for applications by *pwert* scores, and the fit of the functions estimated from equation (1) in section 4 (without controls). Areas of circles are proportional to number of applications that received a particular score. Two models are considered: three discrete jumps (left) and cubic spline (right) as instruments for the nonlinear increase in the funding approval probability. See Section 4 for a description of the models.

Figure 3: Fit for Second Stage Regression: Average propensity to file a European Patent



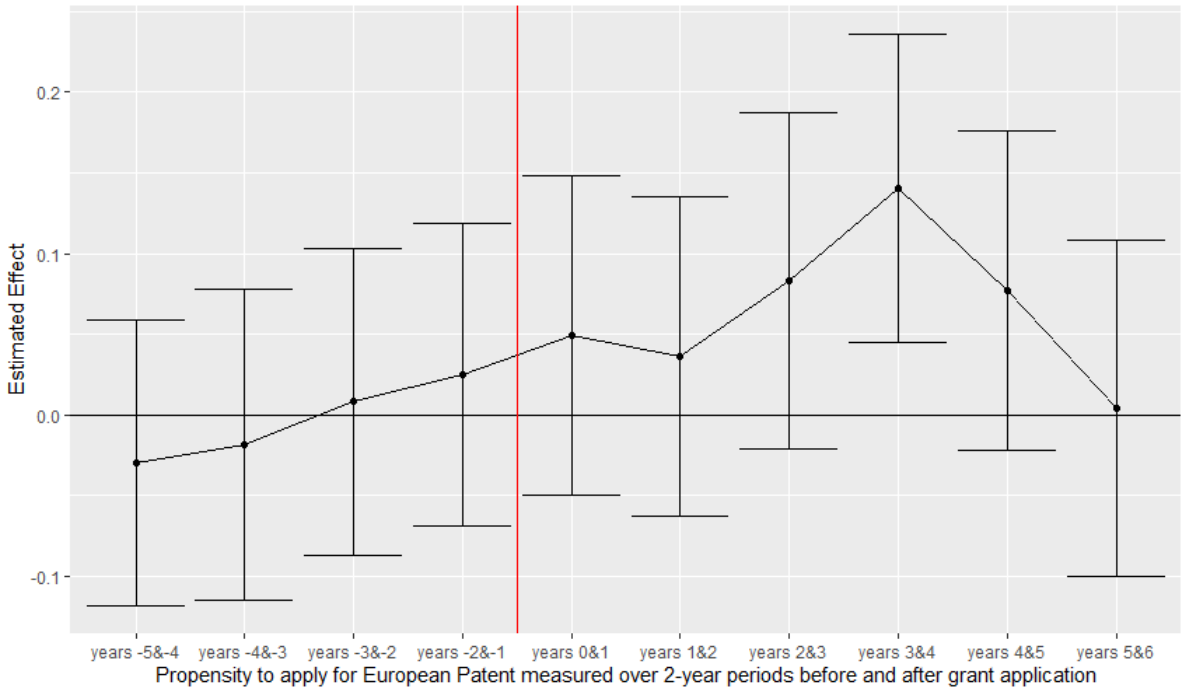
Note: Graphs show the average propensity to file a European Patent for firms by $pwert$ scores, and the fit of the functions estimated from equation (2) in section 4 (without controls). Areas of circles are proportional to number of applications. See section 4 for a discussion. Upper row shows fit for specifications that use three discrete jumps as instruments for the probability of funding approval. Lower row shows fit for specifications that use cubic spline as instrument. Left column shows main specification, where differences in the slope of the direct relationship to the left and to the right of the increase in funding approval are permitted. Right column shows specification that assumes constant slope on entire range. Applications with $pwert$=0 are excluded.

Table 2: The effect of funding approval on the propensity to file a European Patent

	Dependent variable: EP Patent binary post			
	(1)	(2)	(3)	(4)
	Model 1	Model 1	Model 2	Model 2
	IV: 3 jumps	IV: 3 jumps	IV: Cubic Spline	IV: Cubic Spline
approved	0.154**	0.124**	0.128**	0.122**
	(0.0693)	(0.0577)	(0.0634)	(0.0526)
Controls				
Project costs		YES		YES
Prior EP Patents		YES		YES
Prior Non-EP Patents		YES		YES
Firm age group FE		YES		YES
Sector FE		YES		YES
Application year FE		YES		YES
F-statistic (IV)	172.54	161.39	293.20	281.79
Observations	2,480	2,480	2,480	2,480
R-squared	0.045	0.334	0.046	0.334

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, p<0.1. Dependent variable is indicator that takes on value 1 if the firm filed an application for a European Patent in the year of the funding application or during the subsequent three years. Mean of dependent variable is 0.278. Columns 1-2 and 3-4 show results for two alternative instruments for the increase in funding probability, illustrated in Figure 2. F-test statistic for test of joint significance of excluded instruments in the first stage regression shown. F-test statistic in columns 1-2 has three numerator df and in columns 3-4 two numerator df. A discussion of the control variables is included in Section 4.

Figure 4: Time delay of the effect



Note: Figure shows point estimates of the treatment effect and 95% confidence intervals for consecutive two-year periods before and after funding application in percentage points (Year 0 is the year of the funding application), based on the model that uses three discrete jumps as instrument.

Table 3: Heterogenous grant effects by firm age

Dependent variable: EP Patent binary post				
	Young firms		Experienced firms	
	(1)	(2)	(3)	(4)
	Model 1	Model 2	Model 1	Model 2
	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline
Sample	firm age$\leq 5y$	firm age$\leq 5y$	firm age$> 5y$	firm age$> 5y$
approved	0.0421	0.0332	0.210**	0.210***
	(0.0787)	(0.0741)	(0.0828)	(0.0756)
Controls	YES	YES	YES	YES
F-statistic (IV)	62.98	129.13	122.44	163.01
Observations	1,274	1,274	1,206	1,206
R-squared	0.328	0.328	0.336	0.336

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Same controls as in Table 2, except that I exclude firm age group fixed effects. Mean of dependent variable is 0.334 for firms of age$>$ 5y, 0.235 for firms age$\leq$ 5y. Outcomes are measured for the year of funding application and the subsequent three years.

Table 4: The effect of funding approval on the propensity to file a patent by conventionality and by novelty to the firm

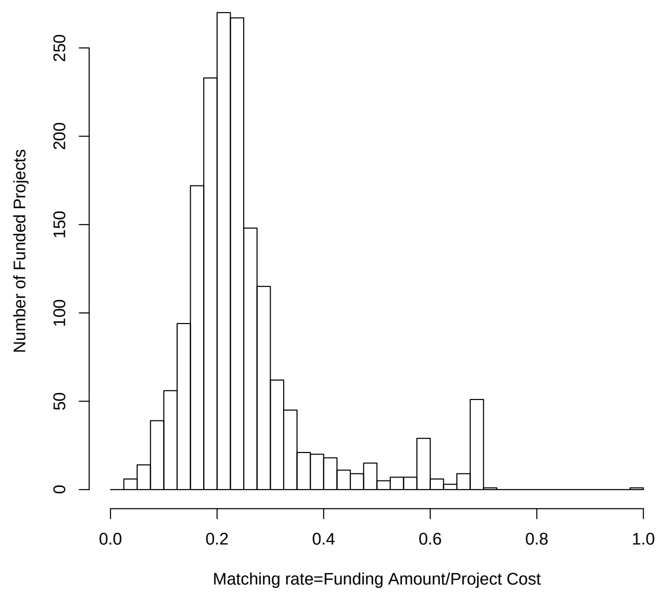
	Bottom 50% Unconventionality		Upper 50% Unconventionality		Top 10% Unconventionality		Does not cite tech. class novel to the firm		Cites tech. class novel to the firm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Model 1	Model 1	Model 1	Model 1	Model 1	Model 1	Model 1	Model 1	Model 1	Model 1
	IV: 3 jumps	IV: 3 jumps	IV: 3 jumps	IV: 3 jumps	IV: 3 jumps	IV: 3 jumps	IV: 3 jumps	IV: 3 jumps	IV: 3 jumps	IV: 3 jumps
Sample	all firms	firm age$>5y$	all firms	firm age$>5y$	all firms	firm age$>5y$	all firms	firm age$>5y$	all firms	firm age$>5y$
							>1 tech. class cited prior to appl.		>1 tech. class cited prior to appl.	
approved	0.0379	0.0337	0.111*	0.192**	0.020	0.103*	0.041	-0.014	0.173	0.206
	(0.0581)	(0.0828)	(0.0618)	(0.0869)	(0.0435)	(0.0609)	(0.116)	(0.135)	(0.124)	(0.137)
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
F-statistic	161.39	122.44	161.39	122.44	161.39	122.44	96.19	169.00	96.19	169.00
Observations	2,480	1,206	2,480	1,206	2,480	1,206	971	571	971	571
R-squared	0.300	0.315	0.340	0.320	0.232	0.278	0.238	0.281	0.172	0.214

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Same controls as in Table 2, except columns 2, 4, 6, 8 and 10 exclude firm age group fixed effects. Columns 1-6 use the patent conventionality measure of Berkes and Gaetani (2020), which considers the frequency of the joint appearance of distinct technological classes (2-digit IPC) among the patent's backwards citations. Columns 7-10 consider whether a patent cites a tech. class (2-digit IPC) that has not been cited by a previous patent of the firm. Mean of dependent variables: 0.253 (column 1), 0.285 (2), 0.296 (3), 0.351 (4), 0.105 (5), 0.141 (6) ,0.529 (7), 0.539 (8), 0.527 (9) and 0.516 (10).

Online Appendices - not for publication

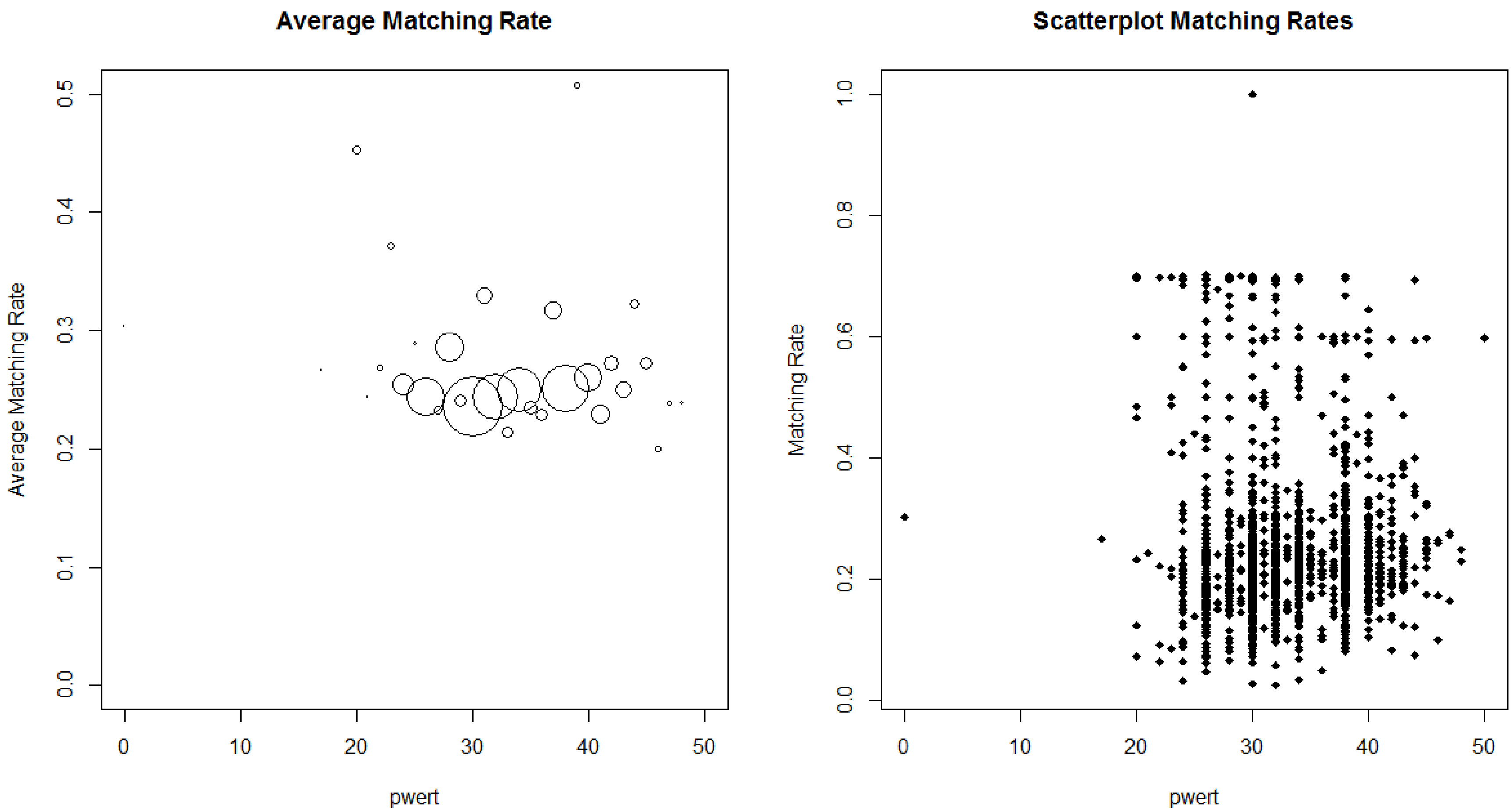
Appendix to section 3

Figure 5: Distribution of matching rates conditional on funding approval in the baseline sample



Note: Matching rate is defined as funding amount divided by total cost of the project.

Figure 6: Matching rates by $pwert$ score conditional on funding approval



Note: Matching rate is defined as funding amount divided by total cost of the project. Left figure shows average matching rates for approved funding applications that received different $pwert$ scores. Areas of circles are proportional to number of approved funding applications that received a particular score. Right figure shows the corresponding scatter plot.

Table 5: Number of funding applications per firm

# of applications by the same firm in the given period	2002	2003	2004	2005	2002-2005
1	763	761	707	691	1419
2	108	137	97	129	424
3	30	26	26	23	200
4	13	6	8	9	109
5	10	4	5	6	54
6	0	4	3	3	38
7	2	1	2	1	20
8	1	0	1	3	14
9	1	0	0	0	14
>=10	3	4	2	2	30

Note: All firms included, not only firms from the baseline sample.

Description of all variables

- **pwert:** score for the evaluation of the technical quality of the project taking integer values between 0 and 50. See section 3.1 for description of how score is obtained and Figure 1 for a depiction of the distribution.

- **Project cost:** cost of the submitted project in Euro as quoted by the firm on the funding application. Distribution presented in Table 1.

- **Approved:** Indicator variable that takes on value 1 if funding appliction was approved. 66 percent of applications in the baseline sample were approved. Around 10 percent of applications were rejected without evaluation ("desk rejects", not included in the baseline sample).

- **EP Patent binary post:** Indicator variable that takes on value 1 if firm filed an application for a (at least one) European Patent in the year of the funding application or during the subsequent three years. A "European Patent" in this paper is a patent family that comprises an application to the European Patent Office. To be clear, for patent applications filed in the same year as the funding application, it is not known whether the patent application was filed before or after the firm applied for funding[17]. Descriptive statistics are presented in Table 1.

  Notes: FFG applicant firms are matched manually to patent applicants with an Austrian country code in PATSTAT by firm name and address. Whenever firms are owned by no more than one private individual, I also match the patents of the owner to the firm. I assign the filing date of earliest patent application in the patent family as the filing date of the European Patent. For around 50 percent of

[17]In section 5, I show that the treatment effect on patent applications in the year of the funding application is negligible and that the effect has considerable delay.

European Patents in my sample, the firm first filed an application for an Austrian patent with the Austrian Patent Office and later filed an application with the European Patent Office.

- **Non-EP Patent binary post:** Indicator variable that takes on value 1 if firm filed an application for a (at least one) Non-EP Patent in the year of the funding application or during the subsequent three years. A Non-EP Patent in this paper is a patent family that does **not** comprise an application to the European Patent Office. Hence, this patent family only comprises applications to National Patent Offices, in most cases only to the Austrian Patent Office. Descriptive statistics are presented in Table 1.

  Notes: FFG applicant firms are matched manually to patent applicants with an Austrian country code in PATSTAT by firm name and address. If the firm is owned by one private individual, I also match the patents assigned to the person. I assign the filing date of earliest patent application in the patent family as the filing date of the European Patent.

- **EP Patent binary pre:** Indicator variable that takes on value 1 if firm filed an application for a (at least one) European Patent in the 12 year preceding the year of the funding application. Descriptive statistics are presented in Table 1.

- **Non-EP Patent binary pre:** Indicator variable that takes on value 1 if firm filed an application for a (at least one) Non-EP Patent in the 12 year preceding the year of the funding application. Descriptive statistics are presented in Table 1.

- **EP Patent pre:** Number of European Patents filed by the firm in the 12 year preceding the year of the funding application. Descriptive statistics are presented in Table 1.

- **Non-EP Patent pre:** Number of Non-EP Patents filed by the firm in the 12 year preceding the year of the funding application. Descriptive statistics are presented in Table 1.

- **Firm sector:** The sector classification used in this paper, along with their corresponding NACE 2 codes and the share of funding applications in my baseline sample:

  – Agriculture and Mining: NACE 2 Codes 01-09, 1.2%

  – Manufacture of Food Products, Oil Products and Wood Products: NACE 2 Codes 10-19, 7 %

  – Manufacture of Metal Products, Electronics and Chemical Products (without Pharmaceuticals): NACE 2 Codes 20,22-29, 41%

  – Research Manufacturing: NACE 2 Code 7210, 1.9%

  – Engineering Services ("Ingenieursbüro" in German): NACE 2 Code 7112, 5.3%

  – Manufacture of Pharmaceutical Products: NACE 2 Code 21, 2%

  – Research Pharmaceuticals: NACE 2 Code 7211, 2.8%

  – Manufacture of Instruments, Sports Goods and other Equipment: NACE 2 Codes 30-34, 17%

  – Power Supply and Construction: NACE 2 Codes 35-42, 3.1%

– Wholesale: NACE 2 Codes 43-46, 4.9%
– Consulting and Financial Services: NACE 2 Code 63, 4.4%
– Software: NACE 2 Code 62, 15.8%
– Rest: 2.6%
– Unassigned: 4.5%

Notes: The firm sector classification in this paper is based on the NACE 2 classification. 80 percent of the firms are classified based on their NACE 2 code in AMADEUS. For firms without AMADEUS match, or that have an uninformative NACE 2 Code in AMADEUS (e.g. "7010 - Activities of Head Office"), I manually impute the sector based on information about the firm on the internet. First, I search the firm on www.firmenabc.at, www.unternehmen24.at and www.moneyhouse.at. If successful, I match the sector of operation mentioned on the site with the corresponding sector in my classification. Otherwise, I search the firm on www.google.com and try to find out about the sector of operation through the firm homepage or other news articles or business service sites that mention the firm. My sector classification is deliberately coarse to reduce the risk of wrong sector assignments during the imputation process. Still, for 4.5 percent of the funding applications, I was not able to determine the firm sector of operation. Such funding applications are assigned to the residual class "Unassigned".

- **Firm age:** defined as year of the funding application minus year of incorporation. Figure 8 and Figure 9 contain the distribution of firm age in the baseline sample. The age of the applicant firm (at the time of the funding application) is available for 2444 applications (93.3 percent of applications) in the baseline sample. For 10 percent of the applications in the sample, the firm age is negative, which means that the firm incorporated after the funding application[18]. The age groups, along with their relative shares, which are included as age group fixed effects in all regressions, are:

  – Younger than 2 years: 31.1%
  – Between 3 and 6 years: 18.0%
  – Between 7 and 15 years: 20.5%
  – Older than 15 years: 30.3%
  – Unknown age: 6.6%

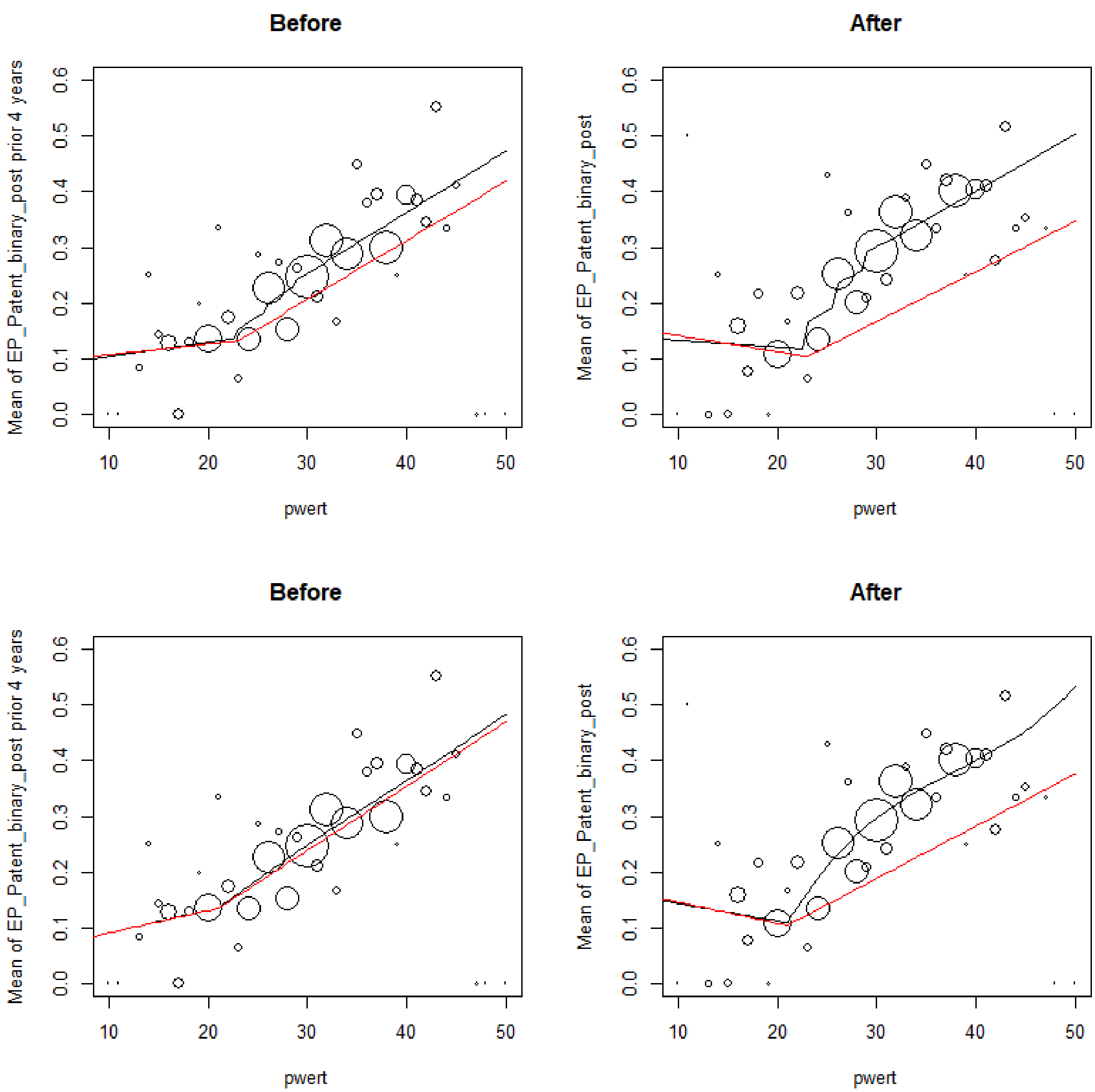
The most plausible reason for why the age may be missing is that the firm never incorporated. For this reason, I regard firms with a missing age as "not established" as of the time of the funding application and assign them to the group of "young" firms in section 5.1. Excluding such firms does not change the findings (not reported).
Notes: For 80 percent of the firms, the year of incorporation is assigned based on the respective entry in AMADEUS. For firms without AMADEUS match, or with an ineligigible entry (year of incorporation is equal 0 or 9999), I manually impute the year of incorporation based on information

[18]The FFG requires grant applicants to incorporate.

about the firm on the internet. First, I search the firm by name and address on www.firmenabc.at, www.unternehmen24.at and www.moneyhouse.at. All of these sources are based on the official Austrian firm register ("Firmenbuch") and registers for one-man businesses that do not meet the sales threshold for the firm register. To the best of my knowledge, the sources track name changes and bankruptcies. There is some noise in the measurement of firm age. If a firm changes its legal form, it reincorporates and resets the official year of incorporation. To the best of my knowledge, there is no way to distinguish genuinely new firms from firms that were formed as a result of a restructuring.

Appendix to section 4

Figure 7: Average propensity of applicant firms to file a European Patent in the 4 years before and in the 4 years after funding application by $pwert$ score



Note: Graphs show the average propensity to file a European Patent in the 4 years before (left column) and after (right column) the funding application, plotted against $pwert$ score. Graph shows fit of the functions estimated from equation (2) in section 4 (without controls), as in Figure 3. Upper row shows fit for model that uses three discrete jumps as instruments for the probability of funding approval. Lower row shows fit for model that uses cubic spline as instrument. See section 4 for detailed description.

Table 6: Placebo Regression (IV Estimates): The effect on the propensity to file a European Patent prior to the the funding application

Dependent variable: EP Patent binary for the 4 years preceding the funding application				
	(1)	(2)	(3)	(4)
	Model 1	Model 1	Model 2	Model 2
	IV: 3 jumps	IV: 3 jumps	IV: Cubic spline	IV: Cubic spline
approved	0.0536	-0.00564	0.0112	-0.0277
	(0.0682)	(0.0559)	(0.0634)	(0.0527)
Controls	NO	YES	NO	YES
Observations	2,480	2,480	2,480	2,480
R-squared	0.035	0.340	0.033	0.339

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Same specifications as in Table 2. Patent control variables are measured 4 years prior to the funding application.

Table 7: Validity check: Applying the IV model that uses 3 jumps as instrument to all control variables

	(1)	(2)	(3)	(4)	(5)	(6)
	project cost	project cost winsorized p=0.95	firm age	firm age winsorized p=0.95	firm age$> 5y$	share manufacturing firms
	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps
approved	-2.079	-0.443	4.514	2.339	0.108	0.122
	(1.900)	(0.686)	(2.874)	(2.209)	(0.0849)	(0.0850)
Controls	NO	NO	NO	NO	NO	NO
Observations	2,480	2,480	2,480	2,480	2,480	2,480
R-squared	0.037	0.070	0.009	0.008	0.007	0.019

	(7)	(8)	(9)	(10)	(11)	(12)
	EP Patents pre	EP Patents pre winsorized p=0.95	EP Patent binary pre	Nat Patents pre	Nat Patents pre winsorized p=0.95	Nat Patent binary pre
	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps	Model 1 IV: 3 jumps
approved	1.782	0.501	0.0261	5.286	0.247	0.0455
	(1.341)	(0.364)	(0.0739)	(3.244)	(0.483)	(0.0747)
Controls	NO	NO	NO	NO	NO	NO
Observations	2,480	2,480	2,480	2,480	2,480	2,480
R-squared	0.009	0.030	0.033	0.002	0.020	0.022

Note: Clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. In columns 2 and 4, the dependent variable is winsorized at the 0.95 percentile. In column 6, the dependent variable is the share of firms in the sectors Manufacture of Food Products, Oil Products and Wood Products, Manufacture of Metal Products, Electronics and Chemicals, Manufacture of Pharmaceutical Products and Manufacture of Instruments, Sports Goods and other Equipment (defined in the Appendix to Section 3).

Table 8: Validity check: Applying the IV model that uses the cubic spline as instrument to all control variables

	(1)	(2)	(3)	(4)	(5)	(6)
	project cost	project cost winsorized p=0.95	firm age	firm age winsorized p=0.95	firm age$> 5y$	share manufacturing firms
	Model 1 IV: cubic spline	Model 1 IV: cubic spline	Model 1 IV: cubic spline	Model 1 IV: cubic spline	Model 1 IV: cubic spline	Model 1 IV: cubic spline
approved	-2.953	-0.861	4.963*	3.040	0.0993	0.127
	(2.049)	(0.665)	(2.664)	(2.011)	(0.0774)	(0.0776)
Controls	NO	NO	NO	NO	NO	NO
Observations	2,480	2,480	2,480	2,480	2,480	2,480
R-squared	0.033	0.068	0.009	0.009	0.009	0.019
	(7)	(8)	(9)	(10)	(11)	(12)
	EP Patents pre	EP Patents pre winsorized p=0.95	EP Patent binary pre	Nat Patents pre	Nat Patents pre winsorized p=0.95	Nat Patent binary pre
	Model 1 IV: cubic spline	Model 1 IV: cubic spline	Model 1 IV: cubic spline	Model 1 IV: cubic spline	Model 1 IV: cubic spline	Model 1 IV: cubic spline
approved	0.944	0.303	0.0085	2.860	0.158	0.0362
	(1.377)	(0.364)	(0.0675)	(1.928)	(0.442)	(0.0679)
Controls	NO	NO	NO	NO	NO	NO
Observations	2,480	2,480	2,480	2,480	2,480	2,480
R-squared	0.012	0.029	0.033	0.002	0.020	0.021

Note: Clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. In columns 2 and 4, the dependent variable is winsorized at the 0.95 percentile. In column 6, the dependent variable is the share of firms in the sectors Manufacture of Food Products, Oil Products and Wood Products, Manufacture of Metal Products, Electronics and Chemicals, Manufacture of Pharmaceutical Products and Manufacture of Instruments, Sports Goods and other Equipment (defined in the Appendix to Section 3).

Appendix to section 5

Table 9: First stage regression for IV models 1-4 in Table 2

Dependent variable: approved, all models estimated with OLS				
	(1)	(2)	(3)	(4)
	Model 1	Model 1	Model 2	Model 2
	3 jumps	3 jumps	Cubic Spline	Cubic Spline
pwert	0.0112***	0.0108***	0.00510	0.00626
	(0.00343)	(0.00378)	(0.00396)	(0.00433)
linear spline pwert 23	-0.00392	-0.00379		
	(0.00381)	(0.00410)		
jump pwert 23	0.3180***	0.3216***		
	(0.0465)	(0.0474)		
jump pwert 26	0.2415***	0.2353***		
	(0.0477)	(0.0478)		
jump pwert 29	0.1690***	0.1647***		
	(0.0306)	(0.0301)		
linear spline pwert 21			0.1558***	0.1508***
			(0.00959)	(0.0102)
quad spline pwert 21			-0.00959***	-0.00931***
			(0.000719)	(0.000735)
cubic spline pwert 21			0.000188***	0.000181***
			(2.02e-05)	(2.06e-05)
project cost (in 100K Euro)		-0.00141		-0.00135
		(0.00140)		(0.00143)
project cost sqr		1.32e-05		1.27e-05
		(9.87e-06)		(1.00e-05)
EP Patent binary pre		0.0182		0.0184
		(0.0187)		(0.0186)
EP Patents pre		-0.00166		-0.00160
		(0.00235)		(0.00235)
EP Patents pre sqr		1.81e-05		2.06e-05
		(2.05e-05)		(2.04e-05)
Non-EP Patent binary pre		0.000245		0.00306
		(0.0180)		(0.0180)
Non-EP Patents pre		0.000135		6.35e-05
		(0.000632)		(0.000639)
Non-EP Patents pre sqr		-1.15e-07		5.30e-08
		(5.50e-07)		(5.55e-07)
Year, Sector, Age Group FE	NO	YES	NO	YES
Observations	2,480	2,480	2,480	2,480
R-squared	0.474	0.495	0.472	0.491

Note: Standard errors in parentheses,*** p<0.01, ** p<0.05, * p<0.1

Table 10: The effect for extended outcome periods

Dependent variable: EP Patent binary post measured for year of the funding application plus subsequent 4, 5 or 6 years						
	Outcomes: 5 years		Outcomes: 6 years		Outcomes: 7 years	
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline
approved	0.154***	0.139***	0.107*	0.094*	0.0884	0.0857
	(0.0590)	(0.0537)	(0.0611)	(0.0561)	(0.0631)	(0.0568)
Controls	YES	YES	YES	YES	YES	YES
Observations	2,480	2,480	2,480	2,480	2,480	2,480
R-squared	0.343	0.345	0.350	0.351	0.350	0.350

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Same specification as in Table 2. Dependent variable is indicator that takes on value 1 if the firm filed an application for a European Patent in the indicated outcome period, starting from the year of the funding application.

Table 11: Robustness check: Alternative specifications

Dependent variable: EP Patent binary post				
Outcomes: 4 years				
	Linear controls		Applications pwert score equal 0 included	
	(1)	(2)	(3)	(4)
	Model 1	Model 2	Model 1	Model 2
	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline
approved	0.095*	0.104**	0.111***	0.108**
	(0.0514)	(0.0521)	(0.0426)	(0.0468)
Controls	YES	YES	YES	YES
Observations	2,480	2,480	2,619	2,619
R-squared	0.336	0.336	0.337	0.337
Outcomes: 5 years				
	Linear controls		Applications pwert score equal 0 included	
	(1)	(2)	(3)	(4)
	Model 1	Model 2	Model 1	Model 2
	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline
approved	0.114**	0.116**	0.129***	0.120**
	(0.0526)	(0.0533)	(0.0435)	(0.0479)
Controls	YES	YES	YES	YES
Observations	2,480	2,480	2,619	2,619
R-squared	0.347	0.346	0.347	0.348

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. In the specification shown in columns 1 and 2, I control for the direct dependence of the propensity to patent on $pwert$ score linearly and exclude the linear spline in the second stage regression, illustrated in Figure 3 (right panel). Second stage equation (2) becomes $EP\ Patent\ binary\ post_i = \alpha' + \tau\ \widehat{approved}_i + \beta_1' pwert_i + \gamma' X_i + u_i$. Specification in columns 3 and 4 includes applications that received $pwert$ equal zero, but is otherwise ident-icalto the main specification. Upper panel shows results when the main outcome variable is measured for the year of the funding appliction plus the subsequent three years (4 years total). Lower panel shows result when thetime frame is extended by one year (5 years total). Same controls included as in Table 2.

Table 12: Robustness check: Extended sample and restricted bandwidth

	Dependent variable: EP Patent binary post						
Outcomes: 4 years							
	Including firms with multiple applications in the same year		Bandwidth		Bandwidth		Using only applications in the range $[19, 22]$ and $[29, 30]$
	Assigned pwert is maximal pwert score of all appl. per firm in given year		15≤pwert≤35		20≤pwert≤30		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	
	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline	IV: pwert$\in [29, 30]$
approved	0.0715	0.0638	0.0712	0.0768	0.104	0.230	0.104***
	(0.0543)	(0.0503)	(0.0917)	(0.0935)	(0.247)	(0.273)	(0.0336)
Controls	YES	YES	YES	YES	YES	YES	YES
Observations	3,103	3,103	1,917	1,917	1,209	1,209	689
R-squared	0.379	0.379	0.341	0.341	0.314	0.292	0.360
Outcomes: 5 years							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	
	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline	IV: pwert$\in [29, 30]$
approved	0.101*	0.0829	0.141	0.130	0.146	0.230	0.126***
	(0.0548)	(0.0506)	(0.0940)	(0.0951)	(0.254)	(0.281)	(0.0346)
Controls	YES	YES	YES	YES	YES	YES	YES
Observations	3,103	3,103	1,917	1,917	1,209	1,209	689
R-squared	0.387	0.387	0.344	0.344	0.320	0.304	0.367

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. In columns 1 and 2, I extend the baseline sample by including firms with multiple funding applications in the same year. Columns 3-6 restrict the baseline sample to applications with a score in the indicated range. Estimated specifications in columns 1-6 as in Table 2. For the model in column 7, first stage equation (1) is given by $approved_i = \alpha + \beta \mathbb{1}_{\{pwert_i \in [29,30]\}}(pwert_i) + \gamma X_i + \epsilon_i$ and second stage equation (2) is given by $EP\ Patent\ binary\ post_i = \alpha' + \tau\ \widehat{approved}_i + \gamma' X_i + u_i$. Note that no controls for the $pwert$ score are included. Upper panel shows results when the main outcome variable is measured for the year of the funding appliction plus the subsequent three years (4 years total). Lower panel shows results when time frame is extended by one year (5 years total).

Table 13: Robustness check: Accounting for the uncertainty about the location of break points and jumps in Figure 2

Dependent variable: EP Patent binary post				
	(1)	(2)	(3)	(4)
	Model 1	Model 1	Model 2	Model 2
	IV: 3 jumps	IV: 3 jumps	IV: Cubic Spline	IV: Cubic Spline
approved				
90% confidence interval	[0.0127;0.2505]	[0.0238;0.2102]	[0.0209;0.2323]	[0.0284;0.2037]
95% confidence interval	[-0.0032;0.2695]	[0.0033;0.2246]	[0.0016;0.2540]	[0.0127;0.2209]
99% confidence interval	[-0.0041;0.3218]	[-0.0259;0.2501]	[-0.0371;0.2918]	[-0.0155;0.2565]
Controls	NO	YES	NO	YES

Note: Bootstrap Procedure: first clusters(=firms) are sampled, location of splines and jumps chosen by best fit from the first stage regression, then the IV model from section 4 is re-estimated. Confidence intervals are reported.

Table 14: Additional results: OLS estimates of the effect on the propensity to file a European Patent

	(1)	(2)	(3)	(4)
	EP Patent binary post	EP Patent binary post	EP Patent binary 4y before	EP Patent binary 4y before
	OLS	OLS	OLS	OLS
			Placebo Test	Placebo Test
approved	0.179***	0.0729***	0.136***	0.0436***
	(0.0187)	(0.0160)	(0.0176)	(0.0151)
Controls	NO	YES	NO	YES
Observations	2,480	2,480	2,480	2,480
R-squared	0.033	0.337	0.021	0.339

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Columns 3 and 4 show the placebo effect of the grant in the 4 years preceding the funding application. In columns 2 and 4, same controls included as in Table 2. In column 4, patent controls are measured 4 years prior to the funding application. OLS estimation equation is given by $EP\ Patent\ binary\ post_i = \alpha + \tau\ approved_i + \gamma X_i + u_i$.

Table 15: Additional results: The effect of funding approval on the number (count) of EP patents

Dependent variable: Number of EP Patents post				
Non-Linear IV Model (Exponential Conditional Mean)				
	Winsorized 99th perc.		Winsorized 95th perc.	
	(1)	(2)	(3)	(4)
	Model 1	Model 2	Model 1	Model 2
	GMM	GMM	GMM	GMM
	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline
approved	0.968	0.896*	0.554*	0.492**
	(0.7302)	(0.5345)	(0.3065)	(0.2306)
Controls	YES	YES	YES	YES
Obs.	2,480	2,480	2,480	2,480
2SLS Linear Model				
	Winsorized 99th perc.		Winsorized 95th perc.	
	(1)	(2)	(3)	(4)
	Model 1	Model 2	Model 1	Model 2
	2SLS	2SLS	2SLS	2SLS
	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline
approved	0.163*	0.157*	0.154*	0.141*
	(0.0985)	(0.0928)	(0.0890)	(0.0841)
Controls	YES	YES	YES	YES
Obs.	2,480	2,480	2,480	2,480
R-squared	0.490	0.490	0.446	0.446

Note: Cluster robust standard errors in parentheses, *** p<0.01,** p<0.05, * p<0.1. Outcome variable is measured for the year of the funding appliction plus the subsequent three years (4 years total). All patent count variables, dependent variable and regressors, are winsorized at the indicated percentile. Upper panel result estimated average marginal effects in the GMM-IV apprach following Windmeijer and Santos Silva (1997). Lower panel shows results for linear 2SLS IV model, where the dependent variable is transformed by the inverse hyperbolic sine. The GMM-IV model assumes an outcome-equation of the form $EP\ Patent\ post_i = \exp(\alpha' + \tau\ \widehat{approved}_i + \beta_1' pwert_i + \beta_2' l_{\geq p*}(pwert_i)$ $+\gamma' X_i + u_i)$. The excluded instrument is given by the predicted propensities obtained from the model $approved_i = \{1 \text{ if } \alpha + \beta_1 pwert_i + \beta_2 l_{\geq p*}(pwert_i) + f_{\geq p*}(\tilde{\beta}; pwert_i)$ $+\gamma X_i + \epsilon_i > 0\}$, where e_i is standard logistic distributed. Standard errors obtained from a clustered bootstrap.

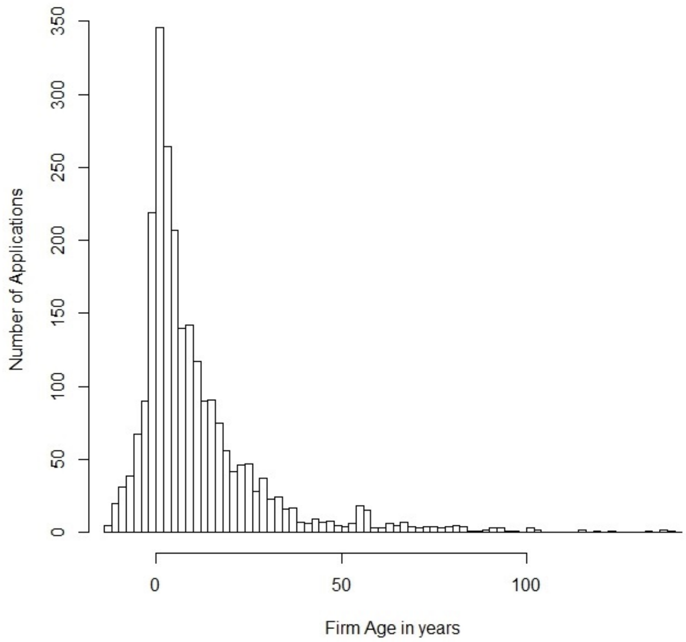
Table 16: Additional results: The effect on the propensity to file a Non-EP Patent

	Dependent variable: Non-EP Patent binary post			
	(1)	(2)	(3)	(4)
	Model 1	Model 1	Model 2	Model 2
	IV: 3 jumps	IV: 3 jumps	IV: Cubic spline	IV: Cubic spline
approved	0.0132	-0.0269	0.0061	-0.0227
	(0.0695)	(0.0606)	(0.0634)	(0.0555)
Controls	NO	YES	NO	YES
Observations	2,480	2,480	2,480	2,480
R-squared	0.029	0.285	0.028	0.286

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Mean of dependent variable is 0.255. Same specification as in Table 2. Dependent variable is indicator that takes on value 1 if the firm filed an application for a Non-EP Patent in the year of funding application or during the subsequent three years.

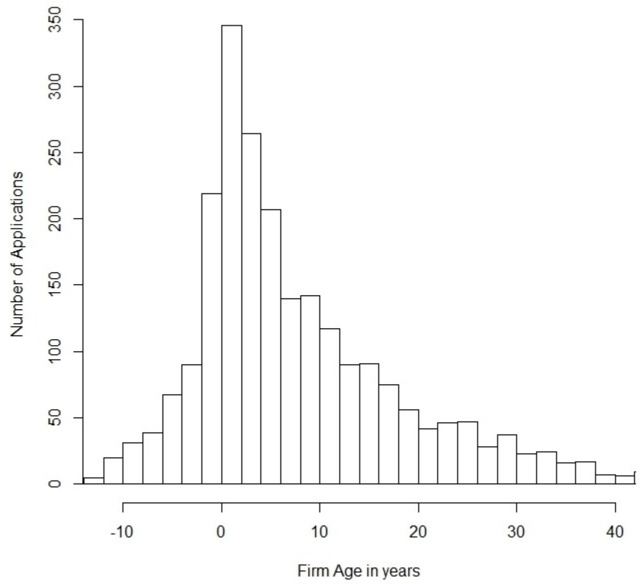
Appendix to section 5.1

Figure 8: Distribution firm age, all firms



Note: Firm age at the time of funding application. A negative firm age indicates that the firm incorporated after the funding application. More details can be found in the Appendix to Section 3.

Figure 9: Distribution firm age, firms of age less than 40 years



Note: Firm age at the time of funding application. A negative firm age indicates that the firm incorporated after the funding application. More details can be found in the Appendix to Section 3.

Table 17: Placebo Regression for heterogeneity in grant effect by firm age

Dependent variable: EP Patent binary for the 4 years preceding the funding application				
	(1)	(2)	(3)	(6)
	Model 1	Model 2	Model 1	Model 2
	IV: 3 jumps	IV: Cubic Spline	IV: 3 jumps	IV: Cubic Spline
	age$\leq 5y$	age$\leq 5y$	age$> 5y$	age$> 5y$
approved	-0.0387	-0.0689	-0.0013	0.0120
	(0.0763)	(0.0754)	(0.0798)	(0.0726)
Controls	YES	YES	YES	YES
Observations	1,274	1,274	1,206	1,206
R-squared	0.287	0.283	0.388	0.388

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Same specification as in Table 6, except firm age group fixed effects are excluded as controls.

Table 18: Robustness check for heterogeneity in grant effect by firm age

Dependent variable: EP Patent binary post

Linear controls

	Young firms		Experienced firms	
	(1) Model 1 IV: 3 jumps	(2) Model 2 IV: Cubic Spline	(3) Model 1 IV: 3 jumps	(4) Model 2 IV: Cubic Spline
	age$\leq 5y$	age$\leq 5y$	age$> 5y$	age$> 5y$
approved	0.0311	0.0287	0.146**	0.166**
	(0.0742)	(0.0742)	(0.0697)	(0.0731)
Controls	YES	YES	YES	YES
Observations	1,274	1,274	1,206	1,206
R-squared	0.327	0.327	0.342	0.341

Applications pwert score equal 0 included

	Young firms		Experienced firms	
	(1) Model 1 IV: 3 jumps	(2) Model 2 IV: Cubic Spline	(3) Model 1 IV: 3 jumps	(4) Model 2 IV: Cubic Spline
	age$\leq 5y$	age$\leq 5y$	age$> 5y$	age$> 5y$
approved	0.0490	0.0363	0.162***	0.170***
	(0.0602)	(0.0659)	(0.0596)	(0.0663)
Controls	YES	YES	YES	YES
Observations	1,361	1,361	1,258	1,258
R-squared	0.322	0.322	0.347	0.346

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Same specifications as in Table 11, except firm age group fixed effects are excluded as controls.

Table 19: Heterogenous treatment effect: firm sales and number of employees vs. firm age

Dependent variable: EP Patent binary post

Subset of applications for which firm sales are available (2144 appl.)

	Heterogeneity by sales				Heterogeneity by age			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.
	sales$\leq 2,797,500$ EUR		sales$> 2,797,500$ EUR		age$\leq 5y$		age$> 5y$	
approved	0.0669	0.0817	0.0813	0.0713	-0.0338	-0.0423	0.210***	0.210***
	(0.0881)	(0.0844)	(0.110)	(0.0916)	(0.0950)	(0.0872)	(0.0915)	(0.0818)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,008	1,008	1,044	1,044	971	971	1,081	1,081
R-squared	0.320	0.320	0.295	0.295	0.337	0.336	0.340	0.340

Subset of applications for which number of employees are available (2144 appl.)

	Heterogeneity by numb. of employees				Heterogeneity by age			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Model 1	Model 3	Model 1	Model 3	Model 1	Model 3	Model 1	Model 3
	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps
	empl≤ 21		empl> 21		age$\leq 5y$		age$> 5y$	
approved	0.111	0.133	0.0102	-0.0023	-0.0338	-0.0423	0.210***	0.210***
	(0.0933)	(0.0904)	(0.106)	(0.0878)	(0.0950)	(0.0872)	(0.0915)	(0.0818)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,007	1,007	1,045	1,045	971	971	1,081	1,081
R-squared	0.292	0.291	0.320	0.320	0.337	0.336	0.340	0.340

Note: Cluster robust standard errors in parentheses, ** p<0.01, ** p<0.05, * p<0.1. Same controls as in Table 3.
Columns 1-4 in the upper panel present split-sample estimates for applications by firms with sales above and below the median amount of sales among applicants. Only applications by firms for which sales are available are considered. Columns 5-8 show split-sample estimates by firm age for this subsample. Analogously, split-sample estimates by number of employees are presented in the lower panel. Firm sales and number of employees are measured in the year prior to the funding application.

Table 20: Heterogenous treatment effect: R&D expenditures and prior patenting experience vs. firm age

Dependent variable: EP Patent binary post

All applications (2619 appl.)

	Heterogeneity by prior patenting experience				Heterogeneity by age			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.
	No prior patent application		At least one prior patent application		age $\leq 5y$		age $> 5y$	
approved	0.0672	0.0762	0.199*	0.168	0.0421	0.0332	0.210**	0.210***
	(0.0636)	(0.0585)	(0.117)	(0.103)	(0.0787)	(0.0741)	(0.0828)	(0.0756)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,528	1,528	952	952	1,274	1,274	1,206	1,206
R-squared	0.104	0.114	0.237	0.238	0.328	0.328	0.336	0.336

Subset of applications where R&D expenditures are available (2046 appl.)

	Heterogeneity by R&D expenditures				Heterogeneity by age			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.	IV: 3 jumps	IV: Cubic Sp.
	R&D Exp $\leq 213,000$ EUR		R&D Exp $> 213,000$ EUR		age $\leq 5y$		age $> 5y$	
approved	0.118	0.0720	0.118	0.121	-0.00814	-0.0272	0.229**	0.218***
	(0.0855)	(0.0854)	(0.126)	(0.104)	(0.100)	(0.0930)	(0.0929)	(0.0828)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	960	960	1,005	1,005	926	926	1,039	1,039
R-squared	0.200	0.201	0.384	0.384	0.340	0.339	0.340	0.342

Note: Cluster robust standard errors in parentheses, ** p<0.01, ** p<0.05, * p<0.1. Same controls as in Table 3.
Columns 1-4 in the upper panel present split-sample for firms that had not filed any patent application (neither at the national nor at the European level) in the 12 years preceding the funding application and firms that had filed at least one patent application during this period. Columns 5-8 present split-sample estimates by firm age, which are reproduced from Table 3. Lower panel shows split-sample estimates for applications by firms with R&D expenditures above and below the median amount of R&D expenditures among applicant firms. R&D expenditures are measured in the year prior to the funding application. Only applications by firms for which R&D expenditures are available are considered. Columns 5-8 in the lower panel show split-sample estimates by firm age for this subsample.

Appendix to section 5.2

Additional results on the utilization of new knowledge: I compute an firm patent measure that captures the average utilization of technological knowledge novel to the firm. This measure, which I call “Average Firm Novelty”, compares the technological classes cited by new patents with the technological classes cited by all previous patents filed by the same firm.[19] For each firm-application pair, I record all distinct technological classes that were cited by any patent filed by the firm in the year of the funding application or during the subsequent three years. Then I count how many of these technological classes had not been cited by any patent filed by the same firm in the pre-application period (between 1980 and the year before the funding application). This number is then normalized by the total number of patents filed in the year of the funding application or during the subsequent three years. For example, for firms applying for funding in 2002, it is defined as follows:

$$avg\_firm\_novelty = \frac{\#\{\text{Novel technological classes cited 2002-2005}\}}{\#\{\text{Patents filed by the firm 2002-2005}\}}$$

The measure has a caveat: it is only meaningful for firm-application pairs with observed patent filings before and after the funding application. The subsample of applications that satisfy these criteria contains 686 funding applications across 485 firms, of which 537 were approved and 149 were rejected. Due to the small sample size, I refrain from using the two-stage model of section 4 and instead compare approved and rejected firms. The results of this descriptive regression can be found in Table 21.

I find that firms that received funding cited on average 0.2 more technological classes per patent that were novel to the firm. However, with a standard error of 0.151, this difference is not statistically significant. Controlling for firm age, firm sector and year of application fixed effects, the estimate of the difference is revised to 0.302 technological classes per patent that were novel to the firm, which is statistically significant at the 0.05 level. The revision is due to the inclusion of sector fixed effects, which suggests that more funding was channeled towards conservative sectors. In the sample, firms cited on average 3.41 distinct technological classes per patent after the grant application, of which 1.16 technological classes were novel to the firm.

Additional information on patent conventionality: “Patent Unconventionality”, introduced by Berkes & Gaetani (2020), measures how “atypical” the combination of technological classes, which are cited by a patent, appear, in the vein of Uzzi et al. (2013). If, for example, a patent cites one patent in IPC class C12 biochemistry and another patent in IPC class A43 footwear, it is considered unconventional because the combination of biochemistry and footwear as knowledge inputs is uncommon. Berkes & Gaetani (2020) compute the technical “relatedness” of IPC classes by calculating the frequency with which two IPC classes were cited by the same patent, using the entire population of patent applications to the USPTO. This symmetric measure for pairs of IPC classes, called “c-score”, is then used to calculate how unconventional the

[19] A ”technological class” corresponds to an IPC class. A patent “cites a technological class” if it cites at least one patent that belongs to the technological class according to the IPC classification.

backward citation structure of any given patent is. I use Berkes & Gaetani's (2020) c-scores and calculate patent unconventionality for a large set of patents by European assignees. I include all patents in PATSTAT filed after 1980 by applicants with a country code from Germany, Italy, Norway, Sweden, Finland, Netherlands, Belgium, Denmark, Austria and Switzerland.[20] Every patent is assigned the minimal c-score across all pairs of distinct IPC classes cited by the patent as the conventionality score, with no further normalization. Then, I rank all patents filed in the same IPC class and year by their conventionality scores and record whether a patent has a conventionality score below or above the median among all patents filed in the same IPC class and year. I classify patents with a below-median score as "Unconventional Patents" and patents with an above-median score as "Conventional Patents". I do not distinguish between National Patents and European Patents. In the sample, for firms of age greater than 5 years, the share of Conventional Patents in all patents filed between 1980 and the year of the funding application, is 47 percent. This share drops to 39 percent in the 4 years after (including the year of) the funding application.

In analogy to the definition of European and National Patents, all patents are in fact DOCDB patent families that possibly comprise multiple patent applications. All citations are computed at the level of DOCDB patent families.

Table 21: Descriptive evidence on research grants and the utilization of knowledge novel to the firm

	(1) avg_firm_novelty_post OLS	(2) avg_firm_novelty_post OLS	(3) avg_firm_novelty_post OLS
approved	0.200	0.302**	0.322**
	(0.151)	(0.139)	(0.137)
Year FE, Sector FE, Age group FE		YES	YES
Other controls			YES
Observations	686	686	686
R-squared	0.003	0.059	0.145

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Dependent variable is the average number of IPC classes cited per patent by the applicant firm that it had never cited prior to the funding application. Outcomes are measured in the year of the funding application and during the three subsequent years. Mean of dependent variable in the sample is 1.16. On average, firms cited 3.41 distinct IPC classes per patent. See section 5.2 for a definition and further discussion.

[20] In many instances, there are not enough Austrian patents by year or technological class to base the ranking just on Austrian patents

Appendix to Section 6

Proof of Proposition 1

The argument is presented after the following Lemma:

Lemma 1. *V_T strictly decreases in T. Furthermore, $V^o < V_T$ for any T.*

Proof. Consider any pair V_T and $V_{\tilde{T}}$ and suppose that $T < \tilde{T}$. Consider the optimal policy when the vintage of the research line is $\tilde{T}$. If the firm followed the same policy starting from a research line of vintage T instead, the incurred payoffs would be strictly higher until the research line would be renewed for the first time (since the stage payoffs of incremental projects are strictly higher for research lines of younger vintage) and exactly equal ever after. Since there is a non-zero chance that the research line is not renewed immediately, this shows that the value of the optimal policy at T must be strictly higher than the value of the optimal policy at $\tilde{T}$. Similarly, consider the optimal policy when there is no existent research line. If the firm followed the same strategy starting from a research line of vintage T instead, the incurred payoffs would be strictly higher until a research line was successfully explored and exactly equal ever after. □

Consider inequality (5). Since V_T and $K\lambda^T$ both strictly decrease in T, and V^0 is lower than V_T for any $T \in \{1, .., N\}$, the right-hand side of (5) strictly increases as vintage T increases. This shows that if exploration is optimal at some T^*, it must also be optimal for all $T > T^*$, where $T \in \{1, .., N\}$. We let $\bar{T}$ be given by the lowest vintage at which exploration is optimal. If exploration is not optimal for any vintage $T \in \{1, .., N\}$, we can set $\bar{T}$ to any number that is strictly greater than N.

Given the assumptions introduced in Section 5, exploration must be optimal if the firm has no research line.

Lemma 2. *A firm with no existent research line undertakes the explorative project.*

Proof. Suppose by contradiction that the optimal strategy is to not undertake the project. Then, the value of a firm with no active research line must be 0. However, the payoff of the (non-stationary) strategy of trying to explore exactly once and then exploit the research line until it becomes obsolete is strictly positive by the assumption that $K + \sum_{t=1}^{N} \beta^t (K\lambda^t - c_i) > c_e/p_e$. This shows that undertaking the project must be optimal for a firm with no active research line. □

Proof of Proposition 2

Part 1: The social planner explores new research lines earlier, i.e. $\bar{T}^* \leq \bar{T}$

The argument is split up into multiple steps. Throughout the proof, I refer to the vintage of the research line as the "state", with states $1, 2, .., N$. For notational simplicity, I denote a firm without an active research line as a firm in state $N + 1$. The state space for firms is therefore given by the natural numbers $\{1, 2, .., N, N + 1\}$. To reiterate, the social planner problem is the firm decision problem with all research payoffs K,$K\lambda$,$K\lambda^2$, .. scaled up by the constant factor $S > 1$.

From Proposition 1 it follows that there are only $N + 1$ candidates for the optimal policy in the firm decision problem and only $N + 1$ candidates for the optimal policy in the social planner problem. The $N + 1$ candidates for the optimal policy are

1. Exploration is optimal in states $T \geq 1$
2. Exploration is optimal in states $T \geq 2$, exploitation is optimal in state $T = 1$
3. Exploration is optimal in states $T \geq 3$, exploitation is optimal in states $T < 3$

and so on. I denote these policies by $\pi^1, \pi^2, \pi^3, .., \pi^{N+1}$ with the indices given by the first state at which exploration happens. Furthermore, I let V_n^S denote the value of policy π^S in state n and denote the value function function of policy π^S by V^S.

Since the state-space is finite, one way to single out the optimal policy is to do the following: first I compute V^{N+1}, the value function of policy π^{N+1}. If the Bellman equations are satisfied in all states $T \geq 1$ for value function V^{N+1}, I stop and conclude that policy π^{N+1} is indeed optimal. If there is any state in which the Bellman equation is not satisfied, I eliminate π^{N+1} as a candidate policy and note that the optimal policy must be in the set $\{\pi^N, \pi^{N-1}, \pi^{N-2}, .., \pi^1\}$. I proceed by computing V^N, the value of policy π^N. Again, if all Bellman equations are satisisfied I stop, otherwise I eliminate π^N and note that the optimal policy must be exploring earlier and lie in $\{\pi^{N-1}, \pi^{N-2}, .., \pi^1\}$. I proceed with candidate policy π^{N-1} and so on.

The argument that proves the statement is the following: in this algorithm, if π^T can be eliminated as a candidate for the optimal policy in the firm decision problem (implying that the firm starts exploring earlier than T), it can also be eliminated as a candidate in the social planner problem. This implies that exploration happens (weakly) earlier in the social planner problem, and for a (weakly) larger number of states. I will state and prove two Lemmas that are jointly sufficient for the result. I repeat the argument after the proof of Lemma 4.

Lemma 3. *If for policy π^T, the Bellman equation is violated in any state $t < T$ in the firm decision problem, it must also be violated in the social planner problem.*

Proof. I start by computing the value of policy π^T in state T, V_T^T, assuming that $T < N + 1$. Policy π^{N+1} will be dealt with at the end of this proof. I denote by s the integer that satisfies $T + s = N$. By definition, the value function V^T satisfies the following set of equations

$$V_t^T = K\lambda^t - c_i + \beta V_{t+1}^T \text{ for all } t \text{ s.t. } 1 \leq t < T \tag{8}$$

and

$$V_{T+n}^T = -c_e + p_e(K + \beta V_1^T) + (1 - p_e)(K\lambda^{T+n} - c_i + \beta V_{T+n+1}^T) \text{ for all } n \text{ s.t. } 0 \leq n \leq s \tag{9}$$

At state $N+1$ it holds that

$$V_{N+1}^T = -c_e + p_e(K + \beta V_1^T) + (1-p_e)(\beta V_{N+1}^T) \tag{10}$$

I now proceed by iteratively plugging equations (9) and (10) into each other to express V_T^T solely as a function of the stage-payoffs, V_{N+1}^T and V_1^T. I obtain

$$\begin{aligned} V_T^T = (1-p_e)^{s+1}\beta^{s+1}V_{N+1}^T + &\sum_{t=0}^{s}(1-p_e)^t\beta^{t+1}p_eV_1^T \\ &- \sum_{t=0}^{s}(1-p_e)^{t+1}\beta^t c_i - \sum_{t=0}^{s}(1-p_e)^t\beta^t c_e + \Psi K \end{aligned} \tag{11}$$

To save on notation, I express the term that multiplies K simply as ΨK. If I make use of the facts that

$$V_{N+1}^T = \frac{p_e(K+\beta V_1^T) - c_e}{1-(1-p_e)\beta}$$

(obtained from equation (10)) and

$$V_1^T = \beta^{T-1}V_T^T + \sum_{t=1}^{T-1}\beta^{t-1}(K\tau^t - c_i)$$

(obtained from iteratively plugging the equations from (8) into each other), I can express V_T^T solely in terms of stage payoffs. (11) becomes

$$\begin{aligned} V_T^T = &\sum_{t=0}^{s}(1-p_e)^t\beta^t p_e(\beta^T V_T^T - \sum_{j=1}^{T-1}\beta^j c_i) + (1-p_e)^{s+1}\beta^{s+1}p_e\frac{\beta^T V_T^T - \sum_{j=1}^{T-1}\beta^j c_i}{1-(1-p_e)\beta} \\ &- \sum_{t=0}^{s}(1-p_e)^{t+1}\beta^t c_i - \sum_{t=0}^{s}(1-p_e)^t\beta^t c_e - \frac{(1-p_e)^{s+1}\beta^{s+1}c_e}{1-(1-p_e)\beta} + \Psi' K \end{aligned} \tag{12}$$

Bringing all V_T^T-terms to the LHS and expressing all geometric sums with the summation formula, this is

$$\begin{aligned} V_T^T\Big\{1 - p_e\beta^T\Big(&\frac{1-\beta^{s+1}(1-p_e)^{s+1}}{1-(1-p_e)\beta} + \frac{\beta^{s+1}(1-p_e)^{s+1}}{1-(1-p_e)\beta}\Big)\Big\} = \\ &- c_i\Big\{\Big(p_e\frac{\beta-\beta^T}{1-\beta}\Big)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta} + \Big(p_e\frac{\beta-\beta^T}{1-\beta}\Big)\frac{(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big\} \\ &- c_i\Big\{(1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big\} \\ &- c_e\Big\{\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta} + \frac{(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big\} + \Psi' K \end{aligned} \tag{13}$$

(13) simplifies to

$$\begin{aligned} &V_T^T\Big(1-\frac{p\beta^T}{1-(1-p_e)\beta}\Big) = \\ &\quad - c_i\Big\{\Big(p_e\frac{\beta-\beta^T}{1-\beta}\Big)\frac{1}{1-(1-p_e)\beta} + (1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big\} \\ &\quad - c_e\frac{1}{1-(1-p_e)\beta} + \Psi' K \end{aligned} \tag{14}$$

With this expression for V_T^T at hand, I can express the condition for exploration for state $t = T-n \leq T-1$ as

$$\begin{aligned} &\frac{c_e}{p_e} - c_i \geq \beta(V_1^T - V_{T-(n-1)}^T) + K - K\lambda^{T-n} \\ &\quad = \beta(\beta^{T-1}V_T^T - \sum_{j=1}^{T-1}\beta^{j-1}c_i - \beta^n V_T^T + \sum_{j=1}^{n}\beta^{j-1}c_i) + \Psi'' K \\ &\quad = -c_i\frac{\beta^{n+1}-\beta^T}{1-\beta} + \frac{\beta(\beta^n-\beta^{T-1})}{1-\frac{p\beta^T}{1-(1-p_e)\beta}}\Big\{\frac{c_e}{1-(1-p_e)\beta} \\ &\quad + c_i\Big(\Big(p_e\frac{\beta-\beta^T}{1-\beta}\Big)\frac{1}{1-(1-p_e)\beta} + (1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big)\Big\} + \Psi''' K \end{aligned} \tag{15}$$

(15) simplifies to

$$\frac{c_e}{p_e} - c_i - \frac{\beta(\beta^n-\beta^{T-1})}{1-\frac{p_e\beta^T}{1-(1-p_e)\beta}}\left\{\frac{c_e-c_i}{1-(1-p_e)\beta} + c_i(1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\right\} \geq \Psi''' K \tag{16}$$

I will now argue that the LHS expression of (16) is in fact strictly positive. Fix $c_e > 0$ and consider the LHS expression of (16) as a linear function of c_i on $[0, c_e/p_e]$.

First, suppose that $c_i = 0$. The LHS expression of (16) is then

$$\begin{aligned} &\frac{c_e}{p_e} - \frac{\beta(\beta^n-\beta^{T-1})}{1-\frac{p_e\beta^T}{1-(1-p_e)\beta}}\frac{c_e}{(1-(1-p_e)\beta)} > \frac{c_e}{p_e} - \frac{\beta(1-\beta^{T-1})}{1-\frac{p_e\beta^T}{1-(1-p_e)\beta}}\frac{c_e}{(1-(1-p_e)\beta)} = \\ &= \frac{c_e}{p_e} - \frac{\beta(1-\beta^{T-1})c_e}{(1-(1-p_e)\beta-p_e\beta^T)} = c_e\frac{1-\beta+p_e\beta-p_e\beta^T-p_e\beta+p_e\beta^T}{p_e(1-(1-p_e)\beta-p_e\beta^T)} = \\ &= c_e\frac{1-\beta}{p_e(1-(1-p_e)\beta-p_e\beta^T)} > 0 \end{aligned}$$

Next, suppose that $c_i = c_e/p_e$. The LHS expression of (16) becomes

$$-\frac{\beta(\beta^n - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}}\left\{\frac{c_e - \frac{c_e}{p_e}}{1-(1-p_e)\beta} + \frac{c_e}{p_e}(1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\right\} =$$
$$= -\frac{\beta(\beta^n - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}}\left\{\frac{c_e - \frac{c_e}{p_e}}{1-(1-p_e)\beta} + \Big(\frac{c_e}{p_e} - c_e\Big)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\right\} =$$
$$= \frac{\beta(\beta^n - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}}\left\{\Big(\frac{c_e}{p_e} - c_e\Big)\frac{(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\right\} \geq 0$$

Since the LHS expression of (16) is linear in c_i, strictly positive at 0 and weakly positive at c_e/p_e, it must be strictly positive for all values in the interior of the interval $[0, c_e/p_e]$.

Now, suppose that condition (16) fails, meaning that the Bellman equation for the firm is not satisified for state $t = T - n$. But then, it must also be true that (16) fails for the social planner in state $T - n$, since, in this case

$$0 < \frac{c_e}{p_e} - c_i - \frac{\beta(\beta^n - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}}\left\{\frac{c_e - c_i}{1-(1-p_e)\beta} + c_i(1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\right\} < \Psi'''K < \Psi'''SK \tag{17}$$

This proves the Lemma for all policies π^T, $T < N+1$. Consider therefore policy π^{N+1}. Repeating all steps from equations (8) to (16), it is straightforward to verify that the value at state $N+1$ for this policy satisfies

$$V_{N+1}^{N+1}\Big(1 - \frac{p_e\beta^{N+1}}{1-(1-p_e)\beta}\Big) = -c_i\frac{p_e\sum_{t=1}^{N}\beta^t}{1-(1-p_e)\beta} - \frac{c_e}{1-(1-p_e)\beta} + \Psi K \tag{18}$$

The analogue to the condition (16) at state $N - n$ is

$$\frac{c_e}{p_e} - c_i - \frac{\beta(\beta^n - \beta^N)}{1 - \frac{p_e\beta^{N+1}}{1-(1-p_e)\beta}}\Big(\frac{c_e - c_i}{1-(1-p_e)\beta}\Big) \geq \Psi'K \tag{19}$$

By the same argument as before, the LHS is strictly positive for all c_i and c_e s.t $c_i < c_e/p_e$. Thus, it must again be the case that if (19) fails in the firm decision problem, it also fails in the social planner problem, since in this case

$$0 < \frac{c_e}{p_e} - c_i - \frac{\beta(\beta^n - \beta^N)}{1 - \frac{p_e\beta^{N+1}}{1-(1-p_e)\beta}}\Big(\frac{c_e - c_i}{1-(1-p_e)\beta}\Big) < \Psi'K < \Psi'SK \tag{20}$$

This concludes the proof. □

Lemma 4. *Suppose that for policy π^T the Bellman equations of the firm decision problem are satisfied in all states $t < T$. Then, at least one policy in $\{\pi^T, \pi^{T+1}, .., \pi^N, \pi^{N+1}\}$ must be optimal in the firm decision* *problem.*

Proof. Since the Bellman equation is satisfied in all states $t < T$, I know that

$$\begin{aligned} V_t^T = K\lambda^t - c_i + &\beta V_{t+1}^T \\ &\geq c_e + p_e(K + \beta V_1^T) + (1 - p_e)(K\lambda^t - c_i + \beta V_{t+1}^T) \text{ for all } t < T \end{aligned} \tag{21}$$

or equivalently

$$\frac{c_e}{p_e} - c_i \geq \beta(V_1^T - V_{t+1}^T) + K - K\lambda^t \text{ for all } t < T \tag{22}$$

I now iterate on the Bellman-operator, starting from value function V^T. Hence, the first iteration is given by

$$V_t^{T,1} = max\{K\lambda^t - c_i + \beta V_{t+1}^T; c_e + p_e(K + \beta V_1^T) + (1 - p_e)(K\lambda^t - c_i + \beta V_{t+1}^T)\} \tag{23}$$

for all t. The subsequent iterations are given by

$$V_t^{T,n} = max\{K\lambda^t - c_i + \beta V_{t+1}^{T,n-1}; c_e + p_e(K + \beta V_1^{T,n-1}) + (1 - p_e)(K\lambda^t - c_i + \beta V_{t+1}^{T,n-1})\} \tag{24}$$

for all t. Because $V^{T,1} \geq V^T$ by construction, it follows from the monotonicity of the Bellman operator that $(V^{T,n})_{n\in\mathbb{N}}$ is an increasing sequence. Since I am iterating on the Bellman-operator, the sequence converges to the value function of the optimal policy. I denote the weakly positive increase in value in iteration n at the fixed state T, $V_T^{T,n} - V_T^{T,n-1}$, by Δ_T^n.

Given $(\Delta_T^n)_{n\in\mathbb{N}}$, I first compute the implied value functions $V_t^{T,n}$ in states $t < T$ and for all subsequent iterations n. I compute the value functions for the first three iterations by hand and then prove the inductive step to establish the validity for all $n \in \mathbb{N}$. For the first iteration, I know that (21) implies that

$$V_t^{T,1} = V_t^T = K\lambda^t - c_i + \beta V_{t+1}^T \text{ for all } t < T \tag{25}$$

Also, by definition

$$V_T^{T,1} = V_T^T + \Delta_T^1 \tag{26}$$

Now, consider the second iteration. First, I have to determine whether exploitation is still optimal in all states $t < T$, given $V^{T,1}$. However, for all states $t < T - 1$

$$\begin{aligned} &\beta(V_1^{T,1} - V_{t+1}^{T,1}) + K - K\lambda^t \\ &= \beta(V_1^T - V_{t+1}^T) + K - K\lambda^t \text{ (by (25))} \\ &\leq \frac{c_e}{p_e} - c_i \text{ (by (22))} \end{aligned}$$

Furthermore, for state $T-1$,

$$\begin{aligned}
&\beta(V_1^{T,1} - V_T^{T,1}) + K - K\lambda^t \\
&= \beta(V_1^T - V_T^T - \Delta_T^1) + K - K\lambda^t \text{ (by (25) and (26))} \\
&\leq \beta(V_1^T - V_t^T) + K - K\lambda^t \\
&\leq \frac{c_e}{p_e} - c_i \text{ (by (27))}
\end{aligned}$$

Thus, exploitation is still optimal in all states $t < T$. This implies that in states $t < T-1$

$$\begin{aligned}
&V_t^{T,2} = \beta V_{t+1}^{T,1} + K\lambda^t - c_i \\
&= \beta V_{t+1}^T + K\lambda^t - c_i \text{ (by (25))} \\
&= V_t^T
\end{aligned}$$

and in state $T-1$

$$\begin{aligned}
&V_{T-1}^{T,2} = \beta V_T^{T,1} + K\lambda^t - c_i = \\
&= \beta(V_T^T + \Delta_T^1) + K\lambda^t - c_i = \\
&= \beta V_T^T + K\lambda^t - c_i + \beta\Delta_T^1 = \\
&= V_{T-1}^T + \beta\Delta_T^1
\end{aligned}$$

I record that $V_t^{T,2} = V_t^T$ for all $t < T-1$ and $V_{T-1}^{T,2} = V_{T-1}^T + \beta\Delta_T^1$. By definition, $V_T^{T,2} = V_T^T + \Delta_T^1 + \Delta_T^2$. It is instructive to also compute the third iteration. Again, I have to determine whether exploration is still optimal in all states $t < T$, given $V_T^{T,2}$. For states $t < T-2$,

$$\begin{aligned}
&\beta(V_1^{T,2} - V_{t+1}^{T,2}) + K - K\lambda^t \\
&= \beta(V_1^T - V_{t+1}^T) + K - K\lambda^t \\
&\leq \frac{c_e}{p_e} - c_i
\end{aligned}$$

For state $T-2$,

$$\begin{aligned}
&\beta(V_1^{T,2} - V_{T-1}^{T,2}) + K - K\lambda^{T-2} \\
&= \beta(V_1^T - V_{T-1}^T - \beta\Delta_T^1) + K - K\lambda^{T-2} \\
&\leq \beta(V_1^T - V_{T-1}^T) + K - K\lambda^{T-2} \\
&\leq \frac{c_e}{p_e} - c_i
\end{aligned}$$

And for state $T-1$,

$$\begin{aligned}
&\beta(V_1^{T,2} - V_T^{T,2}) + K - K\lambda^{T-1} \\
&= \beta(V_1^T - V_T^T - \Delta_T^1 - \Delta_T^2) + K - K\lambda^{T-1} \\
&\leq \beta(V_1^T - V_T^T) + K - K\lambda^{T-2} \\
&\leq \frac{c_e}{p_e} - c_i
\end{aligned}$$

Thus, exploitation is still optimal in all states $t < T$. Thus, I have that

$$V_t^{T,3} = K\lambda^t - c_i + \beta V_{t+1}^{T,2} \text{ for all } t < T \tag{27}$$

Plugging in the expressions for $V_{t+1}^{T,2}$, I record that $V_t^{T,3} = V_t^T$ for $t < T-2$, $V_{T-2}^{T,3} = V_{T-2}^T + \beta^2\Delta_T^1$ and $V_{T-1}^{T,3} = V_{T-1}^T + \beta(\Delta_T^1 + \Delta_T^2)$. Again, by definition $V_T^{T,3} = V_T^T + \Delta_T^1 + \Delta_T^2 + \Delta_T^3$. If I keep expanding, I have that for state $t = T-s$, s.t. $0 \leq s \leq T-1$, in iteration n,

$$V_{T-s}^{T,n} = V_{T-s}^T + \beta^s \sum_{j=1}^{n-s} \Delta_T^j \tag{28}$$

I now prove the inductive step. Suppose (28) holds in iteration n. I will show that (28) holds for iteration $n+1$. First, I check whether in iteration $n+1$, exploitation is still optimal in all states $t < T$ given $V^{T,n}$ and (28). For states $t = T-s$, s.t. $1 \leq s \leq T-1$,

$$\begin{aligned}
&\beta(V_1^{T,n} - V_{T-(s-1)}^{T,n}) + K - K\lambda^{T-s} = \\
&= \beta\Big(V_1^T + \beta^{T-1}\sum_{j=1}^{n-(T-1)}\Delta_T^j - V_{T-(s-1)}^T - \beta^{s-1}\sum_{j=1}^{n-(s-1)}\Delta_T^j\Big) + K - K\lambda^{T-s} = \\
&= \beta\Big(V_1^T - V_{T-(s-1)}^T\Big) + K - K\lambda^{T-s} + \beta\Big(\beta^{T-1}\sum_{j=1}^{n-(T-1)}\Delta_T^j - \beta^{s-1}\sum_{j=1}^{n-(s-1)}\Delta_T^j\Big) \\
&\leq \beta\big(V_1^T - V_{T-(s-1)}^T\big) + K - K\lambda^{T-s} \leq \frac{c_e}{p_e} - c_i
\end{aligned}$$

Hence, exploitation is still optimal in states $t < T$. Thus, for states $t = T-s$, s.t. $1 \leq s \leq T-1$,

$$\begin{aligned}
&V_{T-s}^{T,n+1} = \beta V_{T-(s-1)}^{T,n} + K\lambda^{T-s} - c_i = \\
&= \beta\Big(V_{T-(s-1)}^T + \beta^{s-1}\sum_{j=1}^{n-(s-1)}\Delta_T^j\Big) + K\lambda^{T-s} - c_i = \\
&= V_{T-s}^T + \beta^s \sum_{j=1}^{(n+1)-s}\Delta_T^j
\end{aligned}$$

and by definition

$$V_T^{T,n+1} = V_T^T + \sum_{j=1}^{n+1} \Delta_T^j$$

This establishes (28) for all steps of the iteration. Since $V^{T,n}$ converges to the value function of the optimal policy, call it V, it must be that for states $t = T - s$, s.t. $0 \leq s \leq T - 1$,

$$V_{T-s} = lim_{n \to \infty} V_{T-s}^{T,n} = V_{T-s}^T + \beta^s \sum_{j=1}^{\infty} \Delta_T^j \tag{29}$$

Since $\sum_{j=1}^{\infty} \Delta_T^j$ is just $V_T - V_T^T$, this series is definitely finite. I am now in a position where I can determine whether exploitation is preferred at states $t < T$ under the optimal policy. First, suppose that $\sum_{j=1}^{\infty} \Delta_T^j > 0$. Note that in this case

$$\begin{aligned} &\beta\Big(V_1 - V_T\Big) + K - K\lambda^T \\ &= \beta\Big(V_1^T - V_T^T\Big) + \beta\Big(\beta^{T-1}\sum_{j=1}^{\infty} \Delta_T^j - \sum_{j=1}^{\infty} \Delta_T^j\Big) + K - K\lambda^T \\ &< \beta\Big(V_1^T - V_T^T\Big) + K - K\lambda^T \leq \frac{c_e}{p_e} - c_i \end{aligned}$$

In this case, exploitation is strictly preferred at state $T - 1$ under the optimal policy, which implies that the optimal policy must be an element of $\{\pi^T, \pi^{T+1}, .., \pi^N, \pi^{N+1}\}$. If $\sum_{j=1}^{\infty} \Delta_T^j = 0$, meaning that π^T is optimal, the claim is trivially true. □

Suppose that in the elimination algorithm described in the beginning of the proof, I have already eliminated $\{\pi^{T+1}, .., \pi^N, \pi^{N+1}\}$ as candidates for the optimal policies, for both the firm decision problem and the social planner problem. Hence, the optimal policies of the firm decision problem and the social planner problem must be elements of $\{\pi^1, .., \pi^T\}$. Suppose that all Bellman equations for π^T hold in the firm decision problem. Thus, π^T is the optimal policy in the firm decision problem and the algorithm stops. Since the optimal policy for the social planner is in $\{\pi^1, .., \pi^T\}$, it holds true that the social planner starts exploring weakly earlier than the firm. If at least one Bellman equation for π^T is violated in the firm decision problem, π^T is eliminated as a candidate for the optimal policy in the firm decision problem and the actual optimal policy lies in $\{\pi^1, .., \pi^{T-1}\}$. I will argue that it must be that one of the Bellman equations for the states $t < T$ is violated. By contradiction, suppose not. Then, by Lemma 4, this means that at least one element of $\{\pi^T, \pi^{T+1}, .., \pi^N, \pi^{N+1}\}$ is an optimal policy in the firm decision problem. However, those policies were already eliminated in the earlier rounds. Thus, it must be that at least one of the Bellman equations in some state $t < T$ is violated. Hence, by Lemma 3, this implies that the same Bellman equation is also violated in the social planner problem and π^T can be also eliminated as a candidate for the optimal policy in the social planner problem. This concludes the proof of Part 1.

Part 2: The socially optimal research policy is implemented by offering targeted matching grants to firms *with research lines of vintage T s.t. $\bar{T}^* \leq T < \bar{T}$ for explorative projects with matching rates τ_T^{RG} that solve $\frac{c_e(1-\tau_T^{RG})}{p_e} - c_i = \beta(V_1 - V_{T+1}) + K - K\lambda^T$ where all value functions are obtained from the firm decision problem. The policy minimizes the expected sum of discounted subsidy payments to firms across all policies that implement the socially optimal research policy and that leave the firm at least as good off as in the firm decision problem (participation constraint).*

The grant policy can be inferred by inspection of condition (5). For a firm with a research line of vintage T s.t. $\bar{T}^* \leq T < \bar{T}$ exploration is not privately optimal. This means that (5) is violated. However, as the RHS of (5) is strictly positive (shown in Proposition 1) and c_e/p_e is the only positive term on the LHS, (5) can be made to hold with equality by lowering the term c_e by an appropriately chosen multiplier $(1 - \tau_T^{RG})$.

Note that by making (5) hold with equality, the suggested grant policy keeps the firm indifferent between exploration and exploitation whenever it induces the the firm to undertake the explorative project. This implies that the firm attains the same payoff under the suggested grant policy as in the firm decision problem. By extension, the firm's value function is identical in both problems and given by V, which was derived in the firm decision problem.

Suppose, by contradiction, that there exists another candidate subsidy policy that implements the socially optimal research policy with strictly lower subsidy payments to the firm. The value that the firm attains under this candidate subsidy policy V^c is the expected discounted sum of the research rewards, the research expenses and the subsidy payments received by the agency. Since the candidate subsidy policy implements the socially optimal research policy, the expected research rewards and research expenses must be the same as the rewards and expenses received under the suggested grant policy. The strictly lower subsidy payments to the firm then imply that the value attained under the candidate subsidy policy V^c must be strictly lower than the value attained under the suggested grant policy V. However, this implies that V^c is strictly lower than the value that the firm attains in the firm decision problem (which is also V), thereby violating the participation constraint.

Proof of Proposition 3

Consider conditions (6) and (7). By the same argument as laid out in Proposition 1, Lemma 1, $\tilde{V}_T$ strictly decreases in T and $\tilde{V}^o < \tilde{V}_T$ for any T. Hence, as in Proposition 1, conditions (6) and (7) imply that there exist thresholds $\tilde{T}_{RG}$ and $\tilde{T}_{\neg RG}$, such that exploration is optimal if and only if the vintage of the research line T exceeds $\tilde{T}_{RG}$ when the firm receives a grant, and such that exploration is optimal if and only if T exceeds $\tilde{T}_{\neg RG}$ when it does not receive a grant. Given the assumptions that $\tilde{\tau}^{RG} > 0$ and $D < 1$, the LHS of condition (6) is strictly smaller than the LHS of condition (7). This implies that exploration is always more attractive when the firm receives a grant and hence $\tilde{T}_{RG} \leq \tilde{T}_{\neg RG}$.

By definition of $\tilde{T}_{RG}$, condition (6) is violated for firms with a research line of vintage T s.t. $T < \tilde{T}_{RG}$.

Hence, if the firm receives a grant it undertakes the incremental project, rendering the grant ineffective. By definition of $\tilde{T}_{RG}$ and $\tilde{T}_{\neg RG}$, for vintages T s.t. $\tilde{T}_{RG} \leq T < \bar{T}_{\neg RG}$ condition (6) is satisfied, whereas condition (7) is violated. Hence, the firm only explores when it receives a grant, meaning that the grant is effective. When $T \geq \bar{T}_{\neg RG}$, both conditions are satisfied, implying that the firm explores irrespective of whether or not it receives a grant. This means that the grant is ineffective in a causal sense, as the behavior of the firm is constant across the factual event of receiving a grant and the counterfactual of not receiving a grant.

Proof of Proposition 4

The transition probabilities between different vintages of research lines, as implied by Proposition 3, are listed below. The state space for firms is given by $\{o, 1, ..., N\}$, where o denotes the state where the firm has no existent research line and $T \in \{1, .., N\}$ represents an existent research line of vintage T.

- o: to state 1 with prob. $p = p_e$ and to state o with prob. $p = (1 - p_e)$
- T s.t. $T < \tilde{T}_{RG}$: to state $T + 1$ with prob. $p = 1$
- T s.t. $\tilde{T}_{RG} \leq T < \tilde{T}_{\neg RG}$: to state 1 with prob. $p = p_g p_e$ and to state $T + 1$ with prob. $(1 - p_g p_e)$
- T s.t. $\tilde{T}_{\neg RG} \leq T$: to state 1 with prob. $p = p_e$ and to state $T + 1$ with prob. $(1 - p_e)$
- N: to state 1 with prob. $p = p_e$ and to state o with prob. $(1 - p_e)$

The corresponding Markov transition matrix **M** has the following form:

$$\mathbf{M} = \begin{array}{cc} & \begin{array}{ccccccccccc} o & 1 & 2 & .. & \tilde{T}_{RG} & \tilde{T}_{RG}{+}1 & .. & \tilde{T}_{\neg RG} & \tilde{T}_{\neg RG}{+}1 & .. & N \end{array} \\ \begin{array}{c} o \\ 1 \\ \vdots \\ \tilde{T}_{RG} \\ \vdots \\ \tilde{T}_{\neg RG} \\ \vdots \\ N \end{array} & \begin{bmatrix} (1-p_e) & p_e & 0 & .. & 0 & 0 & .. & 0 & 0 & .. & 0 \\ 0 & 0 & 1 & .. & 0 & 0 & .. & 0 & 0 & .. & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & p_g p_e & 0 & .. & 0 & (1-p_g p_e) & .. & 0 & 0 & .. & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & p_e & 0 & .. & 0 & 0 & .. & 0 & (1-p_e) & .. & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ (1-p_e) & p_e & 0 & .. & 0 & 0 & .. & 0 & 0 & .. & 0 \end{bmatrix} \end{array}$$

Repeating the definitions from Section 6, the firm distribution over different states is $f_0 = (1, 0, .., 0)$ for firms of age cohort 0, and $f_S = f_0 \cdot \mathbf{M}^S$ for firms of age cohort $S > 0$. The share of firms in age cohort S for which the grant is effective is given by $\pi_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = f_S \cdot d(\tilde{T}_{RG}, \tilde{T}_{\neg RG})$, where d is a column vector that takes value 1 in all vintages T s.t. $\tilde{T}_{RG} \leq T < \bar{T}_{\neg RG}$ and 0 otherwise. The average causal effect of the grant on patenting outcome y for firms in age cohort S is given by $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = \pi_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG})\Delta y$. $\Delta y > 0$ is the difference in the patenting outcome if the explorative project is undertaken instead of the incremental project.

Part 1: For firms in age cohort $S < \tilde{T}_{RG}$, $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = 0$.

Note that if $S < \tilde{T}_{RG}$, f_S takes non-zero values only on its first $S+1$ coordinates. Conversely, $d(\tilde{T}_{RG}, \tilde{T}_{\neg RG})$ takes non-zero values only after coordinate $\tilde{T}_{RG}+1$. This implies $\pi_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = f_S \cdot d(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = 0$ for $S < \tilde{T}_{RG}$, and consequently $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = \pi_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG})\Delta y = 0$ for $S < \tilde{T}_{RG}$.

Part 2: For firms in age cohort $S \geq \tilde{T}_{RG}$, there is an effect size $\tau^* > 0$ and a threshold firm age $\bar{S} \geq \tilde{T}_{RG}$ such that $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) \geq \tau^*$ for all $S \geq \bar{S}$.

The transition matrix $\mathbf{M}$ is irreducible and aperiodic (and finite), which implies that it has a unique stationary distribution, denoted f_∞. By definition, f_∞ solves the equation $f_\infty = f_\infty \cdot \mathbf{M}$.

Lemma 5. *f_∞ must be strictly positive on all coordinates.*

Proof. By contradiction, suppose not. Without loss of generality suppose that the coordinate correspondent to state T is zero. However, according to the equation $f_\infty = f_\infty \cdot \mathbf{M}$, since T is entered from state $T-1$ with strictly positive probability, this means that the coordinate correspondent to state $T-1$ must be zero. However, this in turn implies that the coordinate correspondent to $T-2$ must be zero. The argument can be repeated to conclude that the first T coordinates must be zero. Since state o is entered from state N with strictly positive probability, and N is entered from $N-1$ with strictly positive probability and so on, the argument can be further repeated to conclude that all coordinates must be zero. However, the coordinates of the probability vector must sum to 1. $\square$

By the Convergence Theorem for Markov Chains $f_S = f_0 \cdot \mathbf{M}^S \xrightarrow{S\to\infty} f_\infty$ and consequently $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) = \pi_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG})\Delta y \xrightarrow{S\to\infty} f_\infty \cdot d(\tilde{T}_{RG}, \tilde{T}_{\neg RG})\Delta y > 0$. Consider any strictly positive value τ^* that lies below, but arbitrarily close to $f_\infty \cdot d(\tilde{T}_{RG}, \tilde{T}_{\neg RG})\Delta y$. Then, there is a $\bar{S}$ s.t. $\tau_S(\tilde{T}_{RG}, \tilde{T}_{\neg RG}) \geq \tau^*$ for all $S \geq \bar{S}$. This concludes the proof.