Discussion Papers
Statistics Norway
Research department

No. 729
December 2012

Statistics Norway

Statistisk sentralbyrå

Morten S. Henningsen, Torbjørn Hægeland and Jarle Møen

Estimating the additionality of R&D subsidies using proposal evaluation data to control for firms' R&D intentions

Discussion Papers No. 729, December 2012
Statistics Norway, Research Department

Morten S. Henningsen, Torbjørn Hægeland and Jarle Møen

Estimating the additionality of R&D subsidies using proposal evaluation data to control for firms' R&D intentions

Abstract:
Empirical examination of whether R&D subsidies to private firms crowd out private investments has been hampered by problems related to selection. A particular worry is that research intentions and the quality of current research ideas may be correlated with the likelihood of applying for and receiving subsidies. Proposal evaluation data has been put forward as a potential remedy. Using such data from Norway, we do not find strong evidence suggesting that this type of selection creates a severe bias. Proposal evaluation grades strongly predict R&D investments and reduce selection bias in cross-sectional regressions, but there is limited variation in grades within firms over time. This suggests that unobserved project quality is largely absorbed by firm fixed effects. Our best estimate of the short-run additionality of R&D subsidies is 1.15, i.e., a one-unit increase in subsidy increases total R&D expenditure in the recipient firm by somewhat more than a unit. We demonstrate, however, that there is severe measurement error in the subsidy variable. Additionality is therefore likely to be underestimated, and we conclude that measurement errors may be a more important source of bias than selection when panel data are available.

Keywords: Technology policy, R&D subsidies, input additionality, selection, proxy variables

JEL classification: 38, O32, L53, H25, H32

Acknowledgements: Acknowledgement: We have benefited from comments by Tore Ellingsen, Frank Foyn, Carl Gjersem, Svein Olav Nås, Arvid Raknerud and participants at the workshop on R&D Policy Evaluation at the Ministry for Higher Education and Research in Paris in November 2011. The project is financed by the Research Council of Norway.

Address: Jarle Møen, NHH, Department of Finance and Management Science. Hellevn. 30, N 5045 Bergen, Norway. Email: jarle.moen@nhh.no

Torbjørn Hægeland, Statistics Norway, Research Department. E-mail: torbjorn.haegeland@ssb.no

Morten S. Henningsen, Norwegian Ministry of Labour. Email: msh@ad.dep.no

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no
http://ideas.repec.org/s/ssb/dispap.html

For printed Discussion Papers contact:
Statistics Norway
Telephone: +47 62 88 55 00
E-mail: Salg-abonnement@ssb.no

ISSN 0809-733X
Print: Statistics Norway

Sammendrag

Empiriske analyser av hvorvidt FoU-subsidier til private bedrifter fortrenger disse bedriftenes egne FoU-investeringer er beheftet med seleksjonsproblemer. Spesielt kan det være slik at bedriftenes intensjoner om FoU-investeringer og kvaliteten på forskningsideene kan være korrelert med sannsynligheten for å søke og motta støtte. Det har vært foreslått å bruke data fra søknadsevalueringer for å løse dette seleksjonsproblemet. Ved bruk av slike data fra Norge finner vi ingen sterk støtte for en hypotese om at seleksjon av denne typen gir sterk skjevhet i estimatene for effekten av FoU-subsidier på FoU-investeringene. Karakterer fra søknadsevalueringer er en sterk prediktor for FoU-investeringer og reduserer skjevheten i estimater basert på tverrsnittsdata. Det er imidlertid lite variasjon i prosjektkarakterer for samme bedrift over tid, noe som tyder på at prosjektkvalitet i stor grad fanges opp av bedriftsspesifikke effekter. Vårt foretrukne estimat på effekten av FoU-subsidier på bedriftens FoU-investeringer (addisjonaliteten) er 1,15, dvs. at en krone i subsidier medfører litt mer en enn krone økning i FoU-investeringer. Vi viser imidlertid at det er betydelig målefeil i subsidievariabelen i vårt datasett. Dette medfører at addisjonaliteten sannsynligvis er underestimert, og at målefeil kan være en viktigere kilde til skjevhet enn seleksjon når man benytter paneldata.

1. Introduction

The public good nature of innovation and R&D investments has attracted economists' attention for several decades and constitutes the principal justification for subsidies to commercial R&D. There is, however, little consensus in the literature with respect to the effects of such programmes. Do subsidies to commercial R&D crowd out or stimulate private R&D investments? David, Hall and Toole (2000) survey 32 studies and conclude that "the findings overall are ambivalent". Garcia-Quevedo (2004) conducts a meta-analysis of 74 results from 39 studies, and concludes similarly that "the econometric evidence … is ambiguous". Finally, Bronzini and Iachini (2011) list eleven papers published during the last decade and find that only half of them confirm a positive role for public R&D incentives. Given the large amount of public resources devoted to R&D subsidies in all OECD economies, it is important to resolve this ambiguity.

One likely explanation for the lack of clear evidence is econometric problems related to selection, see David, Hall and Toole (2000), Klette, Møen and Griliches (2000), Jaffe (2002) and Cerulli (2010) for surveys. Wallsten (2000) explains the basic problem in one sentence: "Regressing some measure of innovation on the subsidy can establish a correlation between grants and R&D, but it cannot determine whether grants increase firm R&D or whether firms that do more R&D receive more grants." In order to identify the causal effect of R&D subsidies one must answer the following counter-factual question: What would subsidised firms have done if they did not receive subsidies?

Typically, subsidised firms are more R&D intensive than average firms, hence regressing R&D subsidies on private R&D investments will lead to a severe positive bias in the effect estimate. The standard remedy for this selection problem is to include a firm fixed effect or, equivalently, analyse the effect of a *change* in subsidies on the *change* in private R&D investments. This methodology was first introduced by Lichtenberg (1984). Firm fixed effects will pick up aspects such as R&D experience, networks and experience with the application process and technological opportunities in the firm's product group. As pointed out by Klette et al. (2000), however, there may be unobserved transitory effects that invalidates fixed effects estimation in this context. Firms are more likely to apply for subsidies when they have particularly good projects and a particularly good chance of receiving subsidies. At the same time, when firms have particularly good projects one would expect them to undertake more R&D than usual even without subsidies. Kauko (1996) discusses this endogeneity problem in a cross-sectional setting and asserts that "applications for financial support filed by a firm are certainly highly dependent on its intention to invest in R&D."

Finding instrumental variables that solve this endogeneity problem is very challenging, as demonstrated for example by the conflicting results of Lerner (1999) and Wallstein (2000). Both analyse the effects of the SBIR-programme in the US using an IV approach. Reflecting on these difficulties, Jaffe (2002) suggests that evaluation should be built into the design of public research support programmes. More specifically, he notes that in a "canonical research programme" the agency that disburses money for research typically solicits evaluation reports from outside experts and then organises a committee to rank or group the proposals in terms of priority for funding. The agency decides which proposals to fund, based on the available budget, the recommendations of the committee and possibly other criteria not related to proposal quality such as gender, geography and balancing of the grant portfolio e.g. by scientific field. Data generated by such a process can potentially solve the selection problem. To put it simply, Jaffe's idea is to compare projects right above and below the quality cut-off line used by the agency, and also to utilise the randomisation of funding that criteria not related to project quality creates in the data. He suggests using an estimator based on the regression discontinuity design.

Typically, the proposal quality data envisaged by Jaffe is produced, but not recorded systematically and made available to researchers by grant awarding agencies. Although Jaffe's article is well cited, we are only aware of three studies that try to implement his estimation strategy; Jacob and Lefgren (2011), Benavente, Crespi, and Maffioli (2007) and a very recent contribution by Bronzini and Iachini (2011).[1] The first two studies evaluate funding of academic research while Jaffe's main concern was the effectiveness of support to commercial R&D. Bronzini and Iachini (2011) analyse subsidies to commercial R&D, but lack data for the firms' R&D investments. Instead, they rely on balance sheet data and use investments – intangible, tangible and total – as their preferred outcome variable.

In the current paper we explore the value of proposal quality data, gathered by the Research Council of Norway, in estimating the effect of support to industry-led R&D. The Research Council of Norway has for several years emphasised program evaluation, and proposal quality data is available in the PROVIS database established in 1999. We match the PROVIS evaluation data to the Norwegian Business Enterprise R&D statistics that are part of the joint OECD/Eurostat R&D survey, and we discuss different ways of using the available data to identify the causal effect of R&D subsidies on firms' R&D investments. The effect we attempt to identify is the average effect on firms, conditional

[1] A fourth relevant study is Serrano-Velarde (2008) using quantile regressions and regression discontinuity to estimate the impact of R&D subsidies on firm R&D investment under the French ANVAR programme. Serrano-Velarde utilises a discontinuity resulting from program specific eligibility requirements related to form of ownership, rather than proposal evaluation grades. We briefly summarise studies that use the regression discontinuity design in Appendix B.

on the existence and scale of the entire subsidy programme of the Research Council and alternative public sources.

Our contribution is threefold. First, we discuss some practical limitations of the regression discontinuity design suggested by Jaffe (2002). One concern is aggregation from proposal data to annual firm level R&D data when there are many proposals per firm and when the proposed projects last for several years. A more important limitation is that there are many different sub-programmes with different and unclear thresholds for granting subsidies. This problem is related partly to the fact that there are several quality rankings for various aspects of a single proposal, and that it is not entirely clear which one is the most relevant.

Second, because of the difficulties in implementing the regression discontinuity design, we suggest an alternative approach using the variables from the proposal evaluation database as proxies. Proposal-quality grades have previously been used as a proxy variable in the literature on the effect of R&D support on scientific productivity, but to the best of our knowledge not in studies analysing the effect of subsidies to commercial R&D. [2] A particularly interesting variable in this respect is the evaluation panel's expectation of commercial benefit, i.e. the proposed project's contribution to firm profits. This variable can act as a proxy for a firm's incentive to invest in R&D, and therefore also for the intention to invest. When combining this commercial-benefits proxy with unobserved fixed effects, we find, somewhat to our surprise, that the bulk of variation in proposal grades is across firms rather than within firms. This suggests that including firm fixed effects largely solves the selection problem.

Third, comparing data on subsidies reported both by the firms and the granting agency, we document that there is substantial measurement error in the subsidy variable. This will cause a negative bias in the additionality estimates, in particular in specifications that control for unobserved firm fixed effects (cf. Griliches and Hausman, 1986).

In the course of our analysis we present a series of estimates obtained under alternative assumptions and model specifications. Our preferred point estimates suggest that the short-run additionality of subsidies from the Research Council of Norway is positive and in the interval 1.0 to 1.3, i.e., a one-unit increase in the subsidy increases total R&D expenditure in the recipient firm by somewhat more

---

[2] Arora and Gambardella (2005) estimate the effect of grants from the National Science Foundation (NSF) in the USA on impact weighted publications in a five-year window following the grant decision. One of their control variables is the average reviewer score of the proposal, ranging from 1 (excellent) to 5 (very poor). See also Chudnovsky, López, Rossi and Ubfal (2008) for a related analysis.

than one unit. Using a log-log specification, we find that the elasticity of total R&D with respect to subsidies is about 0.20, suggesting that an increase in subsidies gives rise to a moderate increase in total R&D. Hence, there is no evidence that subsidies to commercial R&D crowd out private investments. In light of the many difficulties presented above, our results should be interpreted with caution. Given the extent of measurement errors that we document, however, we believe that our estimates are more likely to be too low than too high.

The rest of this paper is organised as follows: Section 2 discusses further the selection problem and the approach proposed by Jaffe (2002). Section 3 describes the data. Section 4 contains our econometric analysis and Section 5 concludes. We include three appendixes. Appendix A provides an overview of the various dimensions of a research proposal that is assessed as part of the evaluation process in the Research Council of Norway. Appendix B discusses some relevant studies using regression discontinuity. Appendix C reports the results of an attempt to use evaluation grades as instruments rather than proxies.

2. Selection and proposal evaluation data as a potential remedy

The challenge of establishing the counterfactual in the case of governmental support for R&D comes from the fact that recipients of support typically are not a random sample of all possible recipients. Jaffe (2002) discusses this selection issue in detail and considers the following version of the standard model (p. 25 and 31):[3]

(1) $$Y_{it} = \beta_i D_i + \lambda X_{it} + \alpha_i + \mu_t + \omega_{it} + \varepsilon_{it}$$

where Y_{it} is total R&D expenditure of an applicant, or potential applicant, i in year t, and D_i is a dummy variable that is equal to 1 if the applicant has received a grant. X_{it} is a set of firm- and time-varying covariates, and there are four unobservable determinants of research output. First, there is an unobserved firm-specific effect, α_i, and a common unobserved time effect, μ_t. Next, there is a time- and firm specific effect, ω_{it}, which is unobservable by the econometrician, but observable by the granting agency. Finally, there is a genuine error term, ε_{it}, that is assumed to be uncorrelated with X_{it} and D_i. The key challenge is that D_i may be correlated with ω_{it} and α_i because of selection on β_i.

[3] On page 25, Jaffe sets up the equation and discusses the effect of public support on R&D *output*. However, on page 31, he makes it clear that the same selection problems apply when the dependent variable is total R&D expenditure; i.e., when estimating input additionality as we do in this paper. See also his equation (2b) on page 32.

(An agency trying to maximise the impact of its funding will order the projects according to the β_i's and choose as many of the high β_i-projects as possible. This translates into a selection problem because β_i is likely to be correlated with ω_{it} and α_i.) The firm fixed effects can be eliminated using panel data methods, whereas the time-varying unobserved effect, ω_{it}, cannot.

Regression discontinuity design (see, e.g., Imbens and Lemieux, 2008), requires that the granting agency constructs one single variable that sums up the quality of the proposal, and that can be transformed into a unique value for each firm and year. The rate of acceptance should increase in the ranking, jump clearly at one threshold and not jump at other points. The ranking, if incomplete, should have a sufficient range and there must be a sufficient number of data points on either side of the threshold. Furthermore, the method requires that the relationship between the quality ranking and outcome is smooth around the threshold. The regression discontinuity design as proposed by Jaffe implies using a dummy variable for a ranking above the threshold as an instrument for D_i, while conditioning on the quality ranking itself. This will identify the effect of receiving a grant on Y in a "small" region around the threshold.

The necessary data requirements, however, are not always fulfilled. As we will demonstrate below, the proposal quality data from the Research Council of Norway which we believe are typical for similar support programs in other countries, do not seem fully compatible with the regression discontinuity design.[4] For this reason we chose instead to use the quality ranking as a conditioning proxy variable and to control for unobserved firm-specific effects using a standard fixed effects estimator. However, while the regression discontinuity design only requires that we condition on what the granting agency knows, the proxy solution requires that the evaluation data capture all factors that affect both the probability of receiving a subsidy and the R&D investment decisions. It is not obvious that external experts can evaluate the private or social returns to commercial R&D projects with any precision, and this caveat should be kept in mind when interpreting our results.

3. Data and descriptive statistics

Our core data source is the project databases PROVIS and FORISS of the Research Council of Norway. Matching these unique datasets to administrative registers and censuses for firms, such as the

[4] In the future, data that identify which applications were competing directly against each other for funding at a certain point in time, will become available. It is then likely that we will observe a more clear quality cut-off in the grant awarding process and be able to use a regression discontinuity approach.

R&D survey and structural statistics of Statistics Norway, we get a data set that provides detailed information on firms and research support proposals.

3.1. Data sources and key variables

The PROVIS database contains information on every application to the Research Council of Norway for R&D subsidies. The information includes a unique firm identification number, the grades obtained in the evaluation of the proposal with respect to a number of criteria and the project's start and end year. Information on projects that did not receive support is also recorded. Data on the annual amount received as subsidies have been added by the Research Council of Norway from their FORISS database. The PROVIS evaluation scheme was introduced gradually from 1999. We have access to all applications from private-sector firms until 2008.[5] Firms apply for funding from specific thematic programmes within the Research Council of Norway, and each programme is administered by a programme board. The various programmes have different scale (total amount of subsidies). Some programmes exist for a short period, others continue over many years. Applications follow an announcement of available funding within a given programme, and a deadline is set. Applicants for funding from a given programme then compete with each other for a fixed pool of available funds. The frequency of announcements varies over programs, and firms may be eligible under more than one programme.

The programmes analysed seek to promote R&D initiatives in industrial circles and comprise the Research Council of Norway's main instrument for achieving its industry-oriented R&D objectives. The programmes are of the matching grants type, and funding requires at least 50 % co-financing from private enterprise. Formally, the average co-financing is 60-65 %.[6] As pointed out by Klette and Møen (2012), however, it is an open question as to what extent this induces firms to increase their total R&D investments because they may reduce non-subsidised R&D activities upon receiving an R&D grant.

The R&D survey conducted by Statistics Norway is our source of information on R&D investments. The survey has been conducted at the firm level annually since 2001 (every second year before that) and includes all private sector firms with at least 50 employees. Among firms with 10-49 employees, all firms that reported R&D activity in the previous survey are included. Among the remaining firms with 10-49 employees a stratified sample (by industry and size) is used. The 2006 survey also includes

[5] This is industry led R&D or "user directed innovation programmes" (BIP) in the terminology of the Research Council of Norway. See http://www.rcn.no/en/Research_programmes/1184159006970.

[6] The rules for how to calculate the co-financing were perceived to be rather lenient in the period we have data for. The correct number may therefore be closer to 50 %.

a sample of firms with 5-9 employees. The surveys include approximately 4500 firms each year. The R&D statistics include, among other things, information on intramural R&D and R&D subsidies received from various sources. Data on sales are obtained from firm statistics collected separately by Statistics Norway. The use of unique firm identifiers throughout enables us to link data from different sources. We merge firms that are represented in the R&D statistics in at least one of the years between 2001 and 2007 with firms' project proposals from the PROVIS database during the same time period.

Because the R&D survey only covers a subset of firms with less than 50 employees, we are not able to use all applications in the PROVIS database. Table 1A shows the match between the two data sources. There are 13497 firms represented in the R&D surveys in the years 2001 to 2007, but many of these do not conduct R&D. The PROVIS database includes 1480 firms and 4463 applications that have at least one year of proposed activity within the period 2001-2007. There are 631 firms that are represented both in the R&D surveys and in the PROVIS database. These firms have altogether 2444 applications for support, i.e. an average of almost four per firm. 49 % of the firms have only one application, 82 % have at most three, and 96 % have at most 10 applications. Within any given year, between 30 and 36 % of the firms that received subsidies from the Research Council of Norway, received subsidies for more than one project.

In the regression analysis that follows, we want to focus on R&D-performing firms. Hence, we exclude from our sample firms that are never observed with positive R&D. Moreover, since we will estimate regressions with firm fixed effects, we also exclude firms that are observed only once. These firms will not contribute to identification. In Table 1B, we show the match between the remaining firms and the PROVIS database. Of 2570 R&D-performing firms with at least two observations in the R&D surveys, 446 firms have applied for R&D subsidies and are included in the PROVIS database.

Table 1. Match between firms in and the project evaluation database, PROVIS the R&D surveys

Number of firms (number of applications)	A: All firms Firm present in the PROVIS database				B: Firms present in the regression sample Firm present in the PROVIS database			
		Yes	No	All		Yes	No	All
Firm present in R&D surveys	Yes	631 (2444)	12866	13497	Yes	446 (2048)	2124	2570
	No	849 (2019)	-	849	No	-	-	-
	All	1480 (4463)	12866	14346	All	446 (2048)	2124	2570

Firms in the R&D surveys 2001-2007 and firms in the PROVIS database with proposed activity 2001-2007. Part B of the table includes firms in the R&D surveys that are observed at least twice, and at least once with positive R&D.

There are altogether 11368 firm-year observations of the 2570 firms in our regression sample. This implies that there are on average 4.4 observations per firm.

3.2. Aggregating proposal data to the firm level

The PROVIS and FORISS databases are organised with project proposals as the unit of observation, whereas the remaining data, and hence the analyses, are at the firm level. This makes it necessary to aggregate from the proposal level to the firm level. This aggregation concerns two sets of variables, proposal evaluations and the associated project subsidies. Aggregating subsidies is simple. We use the sum of subsidies paid out each year in all projects for each firm.

With respect to evaluations, we want to measure the quality of research ideas, where quality to the firm means the expected present discounted value of future profits generated by money invested in R&D now, and thus the strength of the incentive to invest in R&D. To achieve this we compute the average grade across all proposals for which the proposed period of activity spans the current year.[7] Proposals are evaluated along a number of dimensions, or aspects, resulting in a number of aspect-specific grades. In addition, an overall grade "Total evaluation" is given. The mean of grade *j* is named MGj. We focus on MG11, "Total evaluation", MG5, "Commercial benefit", and MG6, "Relevance and benefit to society". Almost 60 % of the aggregated firm-year level grades are based on only one project proposal. The average number of proposal level grades behind each aggregated firm-year level grade is 2.2. The maximum number of underlying grades is 33.

3.3. Subsidy measures

Our data sources contain two measures of R&D subsidies from the Research Council of Norway. One measure is self-reported subsidies by the firms in the R&D survey (S^R). The other measure comes

[7] We have also used the maximum grade instead of the mean, and this did not change results materially.

from the Research Council's FORISS database (S^C). The FORISS database includes information on realized projects for firms in the PROVIS database that had their proposals accepted. In principle, the two measures should correspond. Table 2 shows that the two measures do not correspond well, in the sense that one is often zero when the other is strictly positive. Of the 1132 observations with $S^C>0$ or $S^R>0$, both are positive in 449 cases. When both are positive, the coefficient of correlation is 0.70.[8] In a bivariate regression, and within the classical measurement error model, this would cause the OLS regression coefficient to be attenuated by 30 % if the two measures are of equal quality.

Table 2. Correspondence between subsidy measures from the R&D surveys (S^R) and the Research Council (S^C)

	$S^R=0$	$S^R>0$	Total
$S^C=0$	10236	278	10514
$S^C>0$	405	449	854
Total	10641	727	11368

The numbers given in the table are firm-year observations. S^R is subsidies reported in the R&D surveys. S^C is subsidies reported in the FORISS database.

We have checked that periodising is not the main reason why the correspondence between the two subsidy measures is poor. Reporting of subsidies at different firm identification numbers is a possible explanation. This error could occur if subsidies are assigned to different firms within a business group in the R&D survey and in the proposal to the Research Council of Norway, or perhaps if firms change identification numbers over time, but that happens only in rare incidents. The most important source of error is probably that the Research Council registers the entire subsidy on the lead firm in projects where two or more firms participate. With respect to self-reported subsidies, errors arise because the person filling in the questionnaire is not accurate enough in finding the exact numbers or in allocating received subsidies to the right governmental agency.[9] [10]

In addition to subsidies from the Research Council of Norway, firms can receive R&D subsidies from EU bodies (S^{EU}) and from Norwegian ministries, Innovation Norway and some other public sources. Innovation Norway is a government agency for the promotion of nationwide industrial development. In the R&D surveys subsidies from ministries, Innovation Norway and others are combined in one

[8] Under the assumptions of the classical errors-in-variables model and if the two measures are of equal quality, this correlation – known as the reliability ratio – measures the fraction of the variance in reported subsidies that is due to true variation in subsidies. See, e.g., Ashenfelter and Krueger (1994) or Bound, Brown and Mathiowetz (2001).

[9] Some measurement errors could probably be avoided by pooling subsidies from different sources, but then we could not estimate the degree of additionality associated with each specific source. The degree of additionality is likely to vary between sources, e.g. because some public financing is given as matching grants subsidies and some as contract R&D.

[10] See Finne (2011) for an assessment of the accuracy of the Norwegian R&D survey.

variable (S^G). It is important to account for subsidies from sources other than the Research Council of Norway, because subsidies may be correlated. Omitting subsidies from alternative sources will then lead to bias in the estimated additionality associated with the subsidies from the Research Council. The presence of measurement error in the subsidy variables implies that our estimates of additionality will be biased towards zero. However, under certain assumptions we may exploit the fact that we have two subsidy measures to obtain better estimates of the true additionality by using an instrumental variables model. We return to this in section 4.5.

Table 3 describes the correspondence between R&D investments and R&D subsidies. About 10 % of the observations with positive R&D investments report positive subsidies. Among observations with zero R&D investments, 66 observations have nonzero subsidies as reported by the Research Council (S^C).

Table 3. R&D investment and subsidy measures from the R&D surveys (S^R) and the Research Council (S^C)

	S^R=0	S^R>0	S^C=0	S^C>0	Total
Intramural R&D =0	3296	0	3230	66	3296
Intramural R&D >0	7345	727	7284	788	8072
Total	10641	727	10514	854	11368

The numbers given in the table are firm-year observations. Intramural R&D and the S^R–subsidy measure are reported in the R&D surveys. The S^C–subsidy measure is reported in the FORISS database.

3.4 Proposal evaluations and R&D investments

The project proposals are evaluated on 11 aspects. Five of these are evaluated by external scientific experts, the remainder by the Research Council of Norway. Aspect 11, ‘Total evaluation’, is evaluated by the Research Council of Norway taking into consideration the external experts’ evaluations. See Table A1 in the appendix for the aspects and a short summary of the assessment criteria. All aspects except ‘Risk’ and ‘Other conditions’ are evaluated on a scale from 1 to 7 with 7 being the top grade. When deciding on the subsidies, the programme board in charge may decide to grant subsidies to many or few of the applicants, they may subsidise only parts of a given project proposal, or for a shorter period than was applied for. This varies across programmes, and may be endogenous to the number and quality of proposals. In most cases, however, the granted subsidy is close to the amount applied for if the proposal receives a subsidy.

Subsidies are only available at certain points in time, and the timing of announcements, proposal deadlines, programme duration and amount of subsidies available can be regarded as exogenous from the firm’s point of view. This introduces random variation in received subsidies, and thus alleviates to some extent the problem of endogeneity of contemporaneous subsidies in an equation that explains intramural R&D.[11]

Table 4 shows the number of proposals and the acceptance rate by grades of Aspect 11, 'Total evaluation' and Aspect 5, 'Commercial benefits', before aggregation to the firm level. Grades are concentrated in the range 3-6, and it is difficult to identify one jump point. Without a clear threshold, and observing that there are *de facto* only four grade levels assigned, it is difficult to apply the approach advocated by Jaffe (2002) to these data. However, the original endogeneity problem arises from an omitted variable. The omitted variable is the potential returns to R&D investments to the firm, and the evaluation data may provide good proxies for this variable.

In the following, we focus on Aspect 5, 'Commercial benefits'. This variable measures the evaluation panel's expectation of the contribution to profits from the proposed project. It should therefore be a useful proxy for the quality of current project ideas in the firm that also affects the decision to invest in R&D. Table 4 shows that the acceptance rate increases with MG5, but of course less strongly than with MG11, because Aspect 5 is only one of several aspects behind Aspect 11.

Table 4. Number of proposals and acceptance rate by grades

	Aspect 11: Total evaluation		Aspect 5: Commercial benefits	
Grade	Number of proposals	Acceptance rate	Number of proposals	Acceptance rate
1	9	0 %	3	33 %
2	41	0 %	16	31 %
3	102	2 %	96	33 %
4	257	35 %	431	49 %
5	510	73 %	426	71 %
6	362	92 %	152	71 %
7	16	100 %	8	100 %
Missing	743	35 %	908	45 %

Proposal level data.

[11] Note, however, that programmes may be anticipated and that the launching of programmes may be correlated with technological opportunities.

Table 5. Variables by grade for Aspect 5: Commercial benefits (MG5)

MG5 (rounded up to nearest integer)	N	Share with Intramural R&D >0	Mean Intramural R&D, if >0	Share with S^R>0	Share with S^C>0	Mean sales (Mill. NOK)
1	2	0.50	135	0.00	0.50	16
2	6	1.00	9996	0.00	0.33	150
3	52	0.79	6666	0.19	0.33	262
4	289	0.91	21763	0.39	0.58	1003
5	514	0.92	41398	0.44	0.78	3453
6	165	0.90	43436	0.41	0.76	1193
7	10	1.00	36747	0.60	1.00	1232
Missing	252	0.89	12126	0.25	0.51	744
Did not apply	10078	0.69	3795	0.02	0.00	361
Total	11368					

All firm year observations. Grade data from proposals have been aggregated to the firm level as explained in section 3.2. S^C is subsidies reported in the FORISS database. S^R is subsidies reported in the R&D surveys.

Table 5 shows how R&D investment, sales and the incidence of subsidies vary by whether the firm has applied for support from the Research Council of Norway, and by the average evaluation of Aspect 5, 'Commercial benefits', for proposals if it did apply. Firms that did not apply are smaller in terms of sales, are less likely to invest in R&D, and invest far less than applicants if they do invest. Among applicants, there is also a clear positive relationship between grade and R&D investment as we would expect: When firms have particularly good research ideas, they invest more in developing them. However, some of this positive correlation may be because of size, as grades tend to increase with average sales. The relationship between grade and whether or not the applicant invests is less strong for average grades above 3, suggesting that when project quality exceeds a certain level, the decision to invest or not is mainly governed by factors less closely related to proposal quality, i.e., grades seem to matter more for the intensive margin than for the extensive margin. Finally, grades aggregated to the firm level are related closely to whether or not the Research Council of Norway grants support. This is as expected although one may question why projects with a top grade on commercial benefit should receive public support. Even in these cases, however, public support is warranted if the technological risk is high or the firms are liquidity constrained.

The fixed effects analyses rely on variation in grades over time within firms. Using only firm-year observations where the firm has applied for support and where MG5 is non-missing, 122 firms have at least two different values of this variable in two different years, and account for 729 valid observations, of which 559 have non-zero MG5. Among the 559 observations, the variance of MG5 cleansed of within-firm means is 0.24, ranging from -1.5 to 2. It is essentially this variation we use to control for the time varying incentive to invest in R&D when we account for firm fixed effects.

Because the fixed effects approach with grades as proxies for research intentions relies on a limited number of firms and limited variation in grades, we will also use a dynamic model where firm fixed effects are replaced by a lagged dependent variable.

3.5. Summary statistics

Table 6 gives summary statistics for key variables. Sales are measured in NOK million, R&D investment and subsidies are measured in NOK 1000, all deflated by the consumer price index to base year 2000. We note that the distributions of sales and intramural R&D are highly skewed, with the means exceeding the medians by a factor of 8 and 4.5, respectively.

Table 6. Descriptive statistics for main variables

	N	Median	Mean	Std. Dev.
Sales	11368	69.3	537.9	5207.8
Intramural R&D	8072	2150.4	9515.8	36273.1
Subsidies from the Research Council in the R&D surveys (S^R)	727	679.6	1456.2	2207.4
Subsidies from the Research Council in the FORISS database (S^C)	854	947.8	1753.1	2485.9
Subsidies from ministries, Innovation Norway and other public agencies (S^G)	728	442.8	3872.3	29355.4
Subsidies from EU bodies (S^{EU})	280	568.7	1654.5	2814.8
Share of S^R in Intramural R&D	727	0.08	0.14	0.18
Evaluation of commercial benefit (MG5)	1038	5	4.647	0.82

Distribution of variables conditional on positive entries. Sales are measured in million real NOK, subsidies and R&D investments are measured in 1000 real NOK. All values are deflated by the consumer price index to base year 2000.

4. Econometric analysis

In the following, we apply two alternative approaches that utilise the available proposal evaluation data to try to estimate the causal effect of subsidies from the Research Council of Norway on intramural R&D in private firms.

4.1. Baseline model

To focus our discussion, we begin with a specification similar to equation (1), but with a continuous rather than a dichotomous subsidy variable.

(2) $$Y_{it} = \beta S_{it} + \lambda X_{it} + \alpha_i + \mu_t + u_{it}$$

The dependent variable, Y_{it}, is intramural R&D of firm i in year t, and S_{it} is the amount of R&D subsidies received from the Research Council of Norway by firm i in year t. Of the two measures we have of this variable, we chose to use the one that is self-reported by the firms (S^R) as this matches the intramural R&D variable and the other subsidy variables in the sense that they all come out of the R&D surveys conducted by Statistics Norway. Other observed variables that affect R&D are contained in X_{it}. These are sales, subsidies from Norwegian ministries, Innovation Norway and other public agencies (S^G), and subsidies from EU bodies (S^{EU}). Time effects, μ_t , capture macroeconomic variations that affect all firms, and firm-specific fixed effects, α_i , capture constant differences in R&D investment between firms over time. Other unobserved factors that influence Y_{it} are captured by the error term, u_{it} .

The parameter of main interest is β, which measures the average effect of subsidies on intramural R&D from the Research Council. If β exceeds unity, there is positive additionality, i.e., one extra unit of subsidy causes firms to invest more than one extra unit in R&D. If β is smaller than one, subsidies partly crowd out private capital, i.e. firms use the subsidy to finance some of the R&D activity that would also have been carried out without the subsidy. A zero coefficient implies full crowding out.

Recall that summary statistics showed substantial variation in sales and intramural R&D within the sample. This translates into a heteroskedasticity problem. It is common that R&D subsidy programmes include firms that vary considerably in scale, but the previous literature is remarkably silent as to whether and how this is handled.[12] In order to reduce heteroskedasticity associated with differences in firm size, we weight the data using a simple method suggested by Park (1966). The weights are determined by first estimating equation (2) without weights and obtaining the coefficient on log sales in a regression of the log of the squared residuals on log sales. Dividing through in equation (2) by sales to the power of this coefficient/2 gives errors that are approximately homoskedastic.[13]

Table 7, columns (1) and (2), show the results of estimating equation (2) in levels with and without firm fixed effects. A matching grants subsidy regime implies a linear relationship between R&D investments and subsidies. We follow e.g. Wallstein (2000) and Lach (2002) and use this as our main

[12] An exception is Lerner (1999) who explicitly notes that the firms were of very different sizes and that a heteroskedasticity problem potentially existed. Because of this, he divides the firms into groups on the basis of sales and calculated heteroskedastic-consistent standard error. Bronzini and Iachini (2011) scale all variables with sales.

[13] Formal tests show that this procedure works very well compared both to no weighting and to scaling all variables with sales. We combine Park's procedure with estimating heteroskedasticity robust standard errors, hence eliminating all heteroskedasticity is not imperative

functional form. The specification without fixed effects includes industry dummies at the two-digit NACE level. In column (1) Park's (1966) procedure implies dividing the equation through by *sales* to the power of 0.16. In the fixed effects regression in column (2), Park's weight is *sales* to the power of 0.20.

In Table 7, column (1), where we do not control for firm fixed effects, we get additionality estimates that are implausibly high. The coefficient associated with subsidies from the Research Council is 4.32. When we control for firm fixed effects, in column (2), the coefficient is reduced to 1.37. The fixed effects estimate implies that when firms receive one unit more in subsidies than they usually do, they do 1.37 unit more R&D. Most likely the high pooled OLS estimate reflects that firms that do more R&D receive more subsidies, i.e. a reversed causality problem. We may also note that the additionality associated with EU grants is very high, 2.77. This is perhaps still implausibly high – but not very precisely estimated. The additionality associated with grants from ministries, Innovation Norway and other public agencies, 0.33, is very low. This is plausible as this includes contract R&D which is not primarily given with the aim to stimulate the firms' own R&D investments.

As an alternative to the linear specification, we apply a log-log model.[14] The survey by David et al. (2000) shows that log-log is also a fairly common functional form in previous studies. Taking logs has the benefit of reducing problems with outliers and heteroskedasticity such that weighting has little effect on the estimates. However, this specification alters the interpretation of the relationship between the variables as the coefficient on log subsidy is an elasticity.

Columns (3) and (4) show estimates of equation (2) in log-log form with and without fixed effects. The estimated elasticitites on subsidies from the Research Council are 0.42 using OLS and 0.22 when introducing firm fixed effects. We see again that the fixed effects estimate is substantially smaller than the pooled OLS estimates, consistent with the former being spuriously high because of omitted firm specific effects. The fixed effect estimate implies that a 1 % increase in the subsidy raises intramural R&D by 0.22 %. For a firm with initial intramural R&D and subsidy equal to the means presented in Table 6, the effect of a marginal increase in the subsidy of 1 % is about NOK 14600, and the implied increase in intramural R&D is about NOK 21900.[15] Hence, a one-unit increase in the subsidy increases

---

[14] The large number of zeros in intramural R&D and subsidies presents a specification problem. We use the approximation that $\ln(z)=0$ if $z=0$, where z is a variable measured in 1000 real NOK.

[15] This number is based on mean intramural R&D for the 727 firms that receive subsidies and can be calculated using the share of subsidies in intramural R&D given in Table 6.

intramural R&D by about 1.5 units. This estimate at the mean is quite close to the fixed effects levels estimate of 1.4 in column (2).

Table 7. Additionality estimates

	Levels form		Log-log form	
	Pooled OLS	Within (FE)	Pooled OLS	Within (FE)
	(1)	(2)	(3)	(4)
S^R	4.323***	1.368***	.416***	.217***
	(1.349)	(.412)	(.0168)	(.0210)
S^{EU}	7.761***	2.774*	.143***	.0604**
	(1.726)	(1.681)	(.0310)	(.0296)
S^G	1.508***	.339***	.360***	.307***
	(.0261)	(.0748)	(.0162)	(.0222)
Sales	4.173***	.544	.372***	.296***
	(.950)	(.707)	(.0366)	(.0712)
Sales squared	-.00722*	.00238		
	(.00376)	(.00213)		
N	11368	11368	11368	11368
Number of firms	2570	2570	2570	2570
R-squared	.348	.038	.212	.062

The dependent variable is intramural R&D. SR is R&D subsidies from the Research Council of Norway. SEU is subsidies from EU bodies. SG is R&D subsidies from Norwegian ministries, Innovation Norway and other public agencies. All specifications include year dummies. Pooled OLS also includes dummies for two-digit NACE group. We correct for heterskedasticity using Park's (1966) procedure in the levels regression. Standard errors allowing for clustering of residuals by firms are reported in parentheses. p < 0.1, ** p < 0.05, *** p < 0.01.

We have done a series of robustness tests on Table 7. First, with respect to correction for heteroskedasticity, we find that the estimated additionality is sensitive to the choice of weights. In Table 7, column (2), the additionality estimate of 1.37 increases to 2.27 if no weights are applied, and it is reduced to 0.94 if we divide through in equation (2) with the square root of sales instead of the optimal Park's weights as explained above. In light of this, it is interesting that correction for heteroskedacticity has received little attention in the previous literature. Second, with respect to sample criteria, the additionality estimate of 1.37 in Table 7, column (2), is reduced to 0.94 if observations with zero R&D are removed, and it is 1.14 if only firms that have applied for subsidies in the year of the observation are included. Finally, if we remove all observations with zero subsidies, we are not able to estimate positive additionality at all from the levels regression, and with the log-log specification in column (4), the additionality estimate is reduced from 0.22 to 0.11, significant at the 5 %-level. Hence, the results depend critically on including firms that do not receive subsidies as a control group.[16]

[16] This is probably the case for most additionality analyses, and is obviously the only source of identification for the many studies that rely on a dummy for whether firms receive subsidies or not.

4.2. Accounting for R&D intentions: A proxy variable approach

A major concern when interpreting additionality estimates obtained from the models presented above, is that subsidies may be endogenous because of correlation with contemporaneous errors, even after eliminating firm fixed effects. We discussed this in Section 1, and in Section 2 it was formalised by having two time-varying error components ω_{it} and ε_{it} in equation (1). The first component, ω_{it}, represents the quality of current research ideas, or the intention to carry out R&D in the absence of subsidies. This intention may be correlated with the likelihood of applying for and receiving subsidies.

We propose to account for ω_{it} by a proxy variable solution, using the mean grade for Aspect 5, 'Commercial benefits' ($\mathrm{MG5}_{\mathrm{it}}$). Recall that $\mathrm{MG5}_{\mathrm{it}}$ is the average of grade 5 over all proposals that potentially spanned the current year for the given firm, as described in the data section. Aspect 5 is meant to measure the net financial gains from completing the project, although it is unclear whether or not this involves conditioning on taking the product to the market, or reaching some other threshold of success (see Appendix A). The proxy solution requires that $\mathrm{MG5}_{\mathrm{it}}$ is redundant in (2) with $u_{it}=\omega_{it}+\varepsilon_{it}$ once ω_{it} is controlled for, and that (X_{it}, S_{it}, μ_t) are uncorrelated with ω_{it}, conditional on $\mathrm{MG5}_{\mathrm{it}}$.[17] Because $\mathrm{MG5}_{\mathrm{it}}$ is missing in some cases, we also include a dummy variable $\mathrm{MG5MISS}_{\mathrm{it}}$ that equals one if the firm applied this year and MG5 is missing. Note that 'applied this year' refers to years spanned by the projects applied for, not the years when proposals were submitted. In addition, a dummy $\mathrm{REJECT}_{\mathrm{it}}$ for 'applied and all proposals rejected' is included, because those who applied have demonstrated that they have an intention to carry out R&D (although perhaps not in the case of rejection).[18] Hence, the reference category is 'did not apply for funding this year'. The estimated equation then becomes

$$(3) \qquad Y_{it} = \beta S_{it} + \lambda X_{it} + \varphi_1 \mathrm{MG5}_{\mathrm{it}} + \varphi_2 \mathrm{MG5MISS}_{\mathrm{it}} + \varphi_3 \mathrm{REJECT}_{\mathrm{it}} + \alpha_i + \mu_t + \tilde{\varepsilon}_{it} .$$

---

[17] This is when the model is estimated with firm fixed effects. For pooled OLS, the firm fixed effects α_i also needs to be uncorrelated with ω_{it} conditional on the proxy.

[18] Kauko (1996) suggests that controlling for applications filed will solve the endogeneity problem. This, however, is only true to the extent that the firms' own evaluation of the R&D projects is not affected by the outcome of the application.

Table 8. Additionality estimates. Levels form. With proposal evaluations

	Pooled OLS	Within (FE) estimator			
	All firms	All firms	All firms	Firms that applied in at least one year	Firms with variation in MG5
	(1)	(2)	(3)	(4)	(5)
S^R	2.756**	1.275***	1.257***	1.458**	1.126
	(1.181)	(.433)	(.434)	(.612)	(.837)
S^{EU}	7.947***	2.727	2.722	3.396	1.769
	(1.78)	(1.692)	(1.693)	(2.173)	(1.775)
S^G	1.486***	.340***	.340***	.326***	.312***
	(.0277)	(.0753)	(.0750)	(.0691)	(.0657)
Sales	3.946***	.539	.542	.757	.788
	(.920)	(.704)	(.704)	(1.698)	(2.073)
Sales squared	-.00628*	.00240	.00240	.00192	.00168
	(.00363)	(.00212)	(.00212)	(.00481)	(.00584)
MG5	2262***	192.8	634.1*	59.73	-64.05
	(440)	(179.7)	(343.4)	(176.2)	(304.5)
MG11			-434.4		
			(277.5)		
MG5MISS	3537*	293.5	1213	307.2	154.1
	(1812)	(873.3)	(947.7)	(776)	(2140)
REJECT	-2125	-701.2	-859	-370.4	650.4
	(1817)	(1165)	(1094)	(1132)	(2357)
N	11368	11368	11368	2208	729
Number of firms	2570	2570	2570	406	122
R-squared	.368	.037	.037	.061	.058

The dependent variable is intramural R&D. S^R is R&D subsidies from the Research Council of Norway. S^{EU} is subsidies from EU bodies. S^G is R&D subsidies from Norwegian ministries, Innovation Norway and other public agencies. MG5 is the evaluation grade on Aspect 5, 'Commercial benefits'. MG11 is the evaluation grade on Aspect 11, 'Total evaluation'. REJECT implies that the firm applied for subsidies, but had the application rejected by the Research Council. All specifications include year dummies. Pooled OLS also includes dummies for two-digit NACE groups. We correct for heterskedasticity using Park's (1966) procedure. Standard errors allowing for clustering of residuals by firms are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8 shows results of estimating the levels model based on equation (3). We find the pooled OLS results in column (1). The coefficient for MG5 is positive and highly significant. A one-unit increase in the commercial-benefits-grade is associated with an increase in intramural R&D of NOK 2.3 million, which is about 24 % for a firm with average R&D. The coefficient for $MG5_{it}$ missing is marginally significant, and corresponds to an average grade around 1.6.[19] This is very low. The rejection dummy is negative, suggesting that rejected applicants tend to invest less in R&D than other

[19] This is seen by dividing 3537 by 2262.

firms, including non-applicants. This is reasonable as they receive a signal that they have low-quality proposals. The additionality estimate associated with subsidies from the Research Council, is 2.75. This is suspiciously high, but much lower than the corresponding estimate in Table 7. Hence, using the grade for commercial benefit as a proxy for unobserved R&D intentions, does seem to reduce the omitted variable bias in pooled OLS.

When including firm fixed effects in column (2), the explanatory power of the proposal data vanishes. This shows that it is the cross-sectional variation that drives the significance in the model without fixed effects. At 1.28, the additionality point estimate is also quite close to the fixed effects estimate without using evaluation data, which was 1.37. Because Aspect 5 may not pick up all relevant information about the firms' R&D intentions, we include in column (3) the mean of grades for Aspect 11, 'Total evaluation'. However, this grade does not seem to add any information about R&D investment, although MG5 now becomes marginally significant. Experimenting with other grades leads to the same conclusion.

Many firms never apply for R&D subsidies, and these are on average small firms with little R&D investments. In Column (4) we limit the sample to firms that applied for subsidies at least once during the data period, in case these have different parameters. Despite the much smaller sample size, the additionality estimate is fairly stable. It increases to 1.46, and naturally looses precision. The grade variables are still insignificant. Finally, in column (5), we restrict the sample further, and only use firms that have applied more than once and have variation in the MG5-variable. These are the firms that contribute directly to identifying the grade coefficient. In this sample of 122 firms the coefficient on $MG5_{it}$ is negative and still insignificant. We also note a drop in estimated additionality. The coefficient is 1.13 and insignificant.

In Table 9, we report the same analyses as in Table 8 for the log-log specification. Introducing the proxy for R&D intensions in pooled OLS reduces the estimated elasticity of intramural R&D with respect to Research Council subsidies from 0.42 to 0.31. This suggests again that evaluation data do reduce endogeneity in cross-sectional regressions. The coefficients on $MG5_{it}$ and $MG5MISS_{it}$ are positive and significant. A one-unit increase in the grade is associated with a 29 % increase in intramural R&D. This is a substantial effect and consistent with the idea that firms with high quality projects will invest more in R&D than others even without subsidies. We may also note that the size of the estimate is similar to what we found in the levels form regression for a firm with average R&D investments. $REJECT_{it}$ is positive and insignificant in this specification. The results in column (2) confirm the level analysis in Table 8. The fixed effects additionality estimates are virtually unaffected by the proxy variable approach, and all the grade variables are insignificant. Dropping firms that never applied to the Research Council has little impact on the estimates, while using only the few firms with variation in MG5 reduces the estimated subsidy elasticity to half. The coefficient of $MG5_{it}$ remains insignificant and close to zero.

We conclude that with the available data and within a fixed effects framework, there is little to gain from adding grades from proposal evaluations.

Table 9. Additionality estimates, models in log-log form. With proposal evaluations

	Pooled OLS	Within (FE) estimator			
	All firms	All firms	All firms	Firms that applied in at least one year	Firms with variation in MG5
	(1)	(2)	(3)	(4)	(5)
ln S^R	.306***	.220***	.221***	.190***	.099***
	(.0217)	(.0215)	(.0215)	(.0252)	(.0323)
ln S^{EU}	.135***	.0675**	.0666**	.0787**	.0544
	(.0294)	(.0294)	(.0294)	(.0342)	(.0423)
ln S^G	.348***	.305***	.305***	.136***	.0505
	(.0158)	(.022)	(.0220)	(.0266)	(.0318)
ln sales	.329***	.301***	.301***	.599***	.590***
	(.0359)	(.0712)	(.0712)	(.132)	(.201)
MG5	.289***	.00344	.0386	.0118	-.0147
	(.0418)	(.0431)	(.0924)	(.0437)	(.0751)
MG11			-.0344		
			(.0771)		
MG5MISS	1.098***	-.423	-.372	-.382	.0704
	(.274)	(.2852)	(.305)	(.284)	(.397)
REJECT	.0741	.271	.257	.188	-.0918
	(.215)	(.205)	(.208)	(.206)	(.226)
N	11368	11368	11368	2208	729
Number of firms	2570	2570	2570	406	122
R-squared	.221	.059	.059	.087	.0819

The dependent variable is ln(intramural R&D). S^R is R&D subsidies from the Research Council of Norway. S^{EU} is subsidies from EU bodies. S^G is R&D subsidies from Norwegian ministries, Innovation Norway and other public agencies. MG5 is the evaluation grade on Aspect 5, 'Commercial benefits'. MG11 is the evaluation grade on Aspect 11, 'Total evaluation'. REJECT implies that the firm applied for subsidies, but had the application rejected by the Research Council. All specifications include year dummies. Pooled OLS also includes dummies for two-digit NACE groups. Standard errors allowing for clustering of residuals by firms are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.3. Accounting for R&D intentions: Using a lagged dependent variable

The models presented above ignore potential adjustment costs and that projects often span two or more calendar years. Adjustment costs imply that past R&D investment affect current investment conditional on current values of the control variables.[20] The problem can be handled by including lagged R&D investments as a regressor, and 12 out of 39 studies used in the meta-analysis by Garcia-Quevedo (2004) use a specification of this type.

As pointed out by Angrist and Pischke (2009, ch. 5.3), it can be challenging to separate fixed effects and lagged dependent variables in applied work. They recommend trying out both specifications to check robustness. The interpretation of the coefficients on the subsidy differs, however, in an important way. With adjustment costs (represented by a lagged dependent variable), the long-run effect of increasing the subsidy exceeds the short-run effect; whereas if the persistence in R&D is attributed to a fixed effect, the impact of increasing the subsidy for one period only lasts one period.

It is plausible that there are both fixed effects and adjustment costs in R&D investments. With fixed effects, however, the lagged dependent variable becomes endogenous, and the best empirical approach would be to use the Arellano-Bond GMM estimator to account for the lagged dependent variable and eliminate the fixed effects. Because this technique makes use of at least two lags of the data, and the within-firm variation in key variables is limited in our data set, we are forced to leave out fixed effects when including lagged intramural R&D. Lagged intramural R&D will account for part of the fixed effects and allows us to exploit the cross-sectional variation in the data, in particular in the proposal evaluations.

Angrist and Pischke (2009) show that if a model erroneously is estimated with fixed effects instead of a lagged dependent variable, the estimated 'treatment effect' of an intervention will be overestimated. However, if the model is estimated using a lagged dependent variable when one should have used a fixed effect, the treatment effect is underestimated. Hence, estimates from a model with a lagged dependent variable and a model with fixed effects may under certain assumptions be seen as the lower and upper bounds of the true treatment effects. In this respect, a dynamic specification complements our fixed effects analysis.

We estimate the following model, where we apply the proxy variable approach combined with lagged R&D as a right-hand side variable:

[20] See, e.g., David et al., 2000, section 2.6, for a discussion.

$$(4) \qquad Y_{it} = \theta Y_{it-1} + \beta S_{it} + \lambda X_{it} + \varphi_1 \text{MG5}_{\text{it}} + \varphi_2 \text{MG5MISS}_{\text{it}} + \varphi_3 \text{REJECT}_{\text{it}} + \alpha_i + \mu_t + \tilde{\varepsilon}_{it}$$

We also estimate a corresponding model for the log-log form. The long run effect of a marginal increase in the subsidy is in this case $\beta/(1-\theta)$.

Table 10. Models with lagged dependent variable

	Levels form	Levels form	Log-log form	Log-log form
	All firms	Firms with variation in MG5	All firms	Firms with variation in MG5
	(1)	(2)	(3)	(4)
Intramural R&D$_{t-1}$	.840***	.737***	.557***	.426***
	(.089)	(.143)	(.0112)	(.0664)
S^R	1.034**	1.562*	.166***	.107***
	(.425)	(.933)	(.0168)	(.0278)
S^{EU}	.604	1.148	.0274	-.0110
	(.746)	(1.515)	(.0194)	(.0210)
S^G	.555***	.637***	.208***	.0235
	(.108)	(.148)	(.0139)	(.0196)
Sales	.301	.600	.213***	.261***
	(.345)	(1.114)	(.0255)	(.0624)
Sales squared	.00202**	.00156		
	(.00102)	(.00354)		
MG5	418.4**	340.5	.0950***	.0681
	(201.8)	(358.5)	(.0274)	(.0536)
MG5MISS	310	-1742	.188	.417
	(999.7)	(3480)	(.186)	(.328)
REJECT	-474.9	-427.5	-.0136	-.277
	(885.8)	(2699)	(.164)	(.182)
N	7793	591	7793	591
Number of firms	2319	120	2319	120
R-squared	.854	.894	.463	.642

The dependent variable is intramural R&D. S^R is R&D subsidies from the Research Council of Norway. S^{EU} is subsidies from EU bodies. S^G is R&D subsidies from Norwegian ministries, Innovation Norway and other public agencies. MG5 is the evaluation grade on Aspect 5, ´Commercial benefits'. MG11 is the evaluation grade on Aspect 11, ´Total evaluation'. REJECT implies that the firm applied for subsidies, but had the application rejected by the Research Council. All specifications include year dummies. Pooled OLS also includes dummies for two-digit NACE groups. We correct for heteroskedasticity using Park's (1966) procedure in the levels regression. Standard errors allowing for clustering of residuals by firms are reported in parentheses. $p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table 10 displays the results from estimating equation (4) in levels and logs. Beginning with the levels regression in columns (1) and (2), lagged intramural R&D has a large coefficient which means that adjustment costs and persistence in R&D investment are large. However, the omitted fixed effects probably inflate the estimate. For the other variables, we see that the results are quite similar to the fixed effects estimates in Table 8. The estimated additionality is 1.03 in column (1) and 1.56 in column (2). The small subsample of firms with variation in $MG5_{it}$ now gives the largest additionality estimate, while in Table 8 this sample produced the smallest estimate. This instability may simply be due to the small sample size and low precision.

The implied long run effects are very large, 6.5 and 5.9, respectively. Given that we most likely have an omitted fixed effect that is correlated with the included lagged dependent variable, however, we should not put too much emphasis on the long-run effects.

Following the reasoning of Angrist and Pischke (2009) and combining the estimated contemporaneous effect, β_2 , from column (1) in this model and the corresponding fixed effects estimates in Table 8, column (2), suggests that the true short-run additionality effect is within the interval 1.030 to 1.275 with a mid-point of 1.15.

With the log-log specification, the estimates in Table 10, columns (3) and (4), also reveal considerable persistence in R&D investments. The contemporaneous elasticities are fairly close to the fixed effects estimates. For all firms, the interval suggested by Angrist and Pischke (2009) is 0.17 to 0.22. The estimated long-run elasticities range from 0.19 to 0.37. Interestingly, MG5 is significant in column (1) and (3), but the significance disappears when the sample is restricted to firms that have variation in MG5. This suggests that the effect of MG5 in columns (1) and (3) is driven by between-firm variation.

4.5. Measurement error bias

In Section 3.3, we demonstrated that the subsidy measures contain measurement error. When both the available measures are positive, the correlation between them is only 0.70.

Because we have two measures of the same subsidy, S^R and S^C, we may use one as an instrument for the other to reduce measurement error bias, provided that they are only correlated through the true subsidy (see, e.g., Ashenfelter and Krueger, 1994 or Bound, Brown and Mathiowetz, 2001). Because the two observed measures tend to coincide on the value zero (we have many more observations with a zero subsidy than with a positive subsidy), the two measurement errors are correlated.[21] However, this is unconditional on the included variables. By accounting for whether or not the firm has applied

[21] Another situation that will take us outside the classical measurement error model is if total subsidies are measured correctly, but distributed erroneously on the three subsidy variables included in our regression.

for support in a given year through the grade variables and for rejection, we eliminate some of this correlation. We apply the IV solution, using S^C as an instrument for S^R, to both the levels and log-log forms, with and without fixed effects, see Tables 11 and 12. For comparison we include the results from the corresponding regressions in Tables 8 and 9 where we did not attempt to correct for measurement errors.

Table 11. IV estimation to correct for measurement error. Levels form

	Pooled OLS	Pooled OLS	Within (FE)	Within (FE)
Instrument:	None	S^C	None	S^C
	(1)	(2)	(3)	(4)
S^R (endogenous)	2.756**	11.100***	1.275***	2.591*
	(1.181)	(3.493)	(.433)	(1.502)
S^{EU}	7.947***	4.051	2.727	2.523
	(1.780)	(3.006)	(1.692)	(1.642)
S^G	1.486***	1.463***	.340***	.336***
	(.0277)	(.056)	(.0753)	(.0729)
Sales	3.946***	3.835***	.539	.526
	(.920)	(.863)	(.704)	(.702)
Sales squared	-.00628*	-.00642*	.00240	.00236
	(.00363)	(.00339)	(.00212)	(.00211)
MG5	2262***	803.9	192.8	79.4
	(440)	(568.5)	(179.7)	(256.7)
MG5MISS	3537*	-628	293.5	425.9
	(1812)	(2315)	(873.3)	(893.4)
REJECT	-2125	3582	-701.2	-286.7
	(1817)	(2290)	(1165)	(1384)
Coefficients on instrument in first stage:				
S^C		.468***		.502***
		(.0784)		(.0928)
N	11368	11368	11368	11368
Number of firms	2570	2570	2570	2570
R-squared	.368	.312	.0367	.069

The dependent variable is intramural R&D. All specifications include year dummies. Pooled OLS also includes dummies for twodigit NACE groups. Weighted using Park's optimal weight. Standard errors allowing for clustering of residuals by firm are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.0$

Using the levels form, we see that the additionality estimates increase by a factor of about four when using pooled OLS, and by a factor of about two when using fixed effects. The direction is as expected, but the magnitude is too large to be plausible, and suggests that there are unresolved biases. Using the log-log specification, we get about a 30 % increase in the coefficient when using IV on the specification without fixed effects, but in the fixed effects specification the coefficient becomes small and insignificant when using IV. We have no explanation for this, but given the large standard error of the coefficient, it might be coincidental.

In general, we must conclude that our attempt to use IV estimation to correct for measurement errors has limited success. Most likely, this is because the assumptions behind the classical errors-in-variables model are not fulfilled. We have experimented with several alternative specifications and subsamples without obtaining any robust findings. Despite this, something has still been learned from the two measures of subsidies, as we have demonstrated that there are severe measurement errors in the data. Statistics Norway collect business enterprise R&D data following the same procedures as the rest of the OECD, and it is therefore no reason to think that the data quality is much different in other countries. This implies that additionality estimates that do not correct for measurement errors are likely to have a large negative bias.

Table 12. IV estimation to correct for measurement error. Log-log form

	Pooled OLS		Within (FE)	
Instrument:	None	ln S^C	None	ln S^C
	(1)	(2)	(3)	(4)
ln S^R (endogenous)	.306***	.397***	.220***	.117
	(.0217)	(.115)	(.0215)	(.142)
ln S^{EU}	.1349***	.103**	.0675**	.0853*
	(.0294)	(.0478)	(.0294)	(.0493)
ln S^G	.348***	.340***	.305***	.296***
	(.0158)	(.0192)	(.022)	(.0240)
ln sales	.329***	.326***	.301***	.266***
	(.0359)	(.0360)	(.0712)	(.0701)
MG5	.289***	.231***	.00344	.0314
	(.0418)	(.0816)	(.0431)	(.0728)
MG5MISS	1.098***	.890**	-.423	-.365
	(.274)	(.354)	(.285)	(.322)
REJECT	.0741	.298	.271	.125
	(.215)	(.346)	(.205)	(.303)
Coefficients on instrument in first stage:				
ln S^C		.5996***		.500***
		(.0678)		(.0895)
N	11368	11368	11368	11368
Number of firms	2570	2570	2570	2570
R-squared	.221	.220	.059	.069

The dependent variable is ln(intramural R&D). All specifications include year dummies. Pooled OLS also includes dummies for two-digit NACE groups. Standard errors allowing for clustering of residuals by firms are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5. Conclusion

Empirical examination of whether R&D subsidies to private firms crowd out private investments has been hampered by problems related to selection bias. In particular, subsidies may be endogenous even after eliminating firm fixed effects because the quality of current research ideas may be correlated with the likelihood of applying for and receiving subsidies. Access to proposal evaluation data has been suggested as a potential remedy. Using such data we do not find strong evidence suggesting that this type of selection creates a severe bias. Proposal evaluation grades from the Research Council of Norway strongly predict R&D investments in cross-sectional regressions, with a one-unit increase in grade being associated with an increase in intramural R&D between 20 and 30 %. There is, however, limited variation in the proposal evaluation grades within firms over time. Evaluation grades are by no means a perfect measure of project quality, but our findings suggests that unobserved project quality is largely absorbed by firm fixed effects. Selection bias is therefore not likely to be a major issue when panel data are available.

Our best estimate of the short-run additionality of R&D subsidies from the Research Council of Norway is 1.15, i.e., a one-unit increase in the subsidy increases total R&D expenditure in the recipient firm by somewhat more than one unit. This is slightly higher than the fixed effects estimate in Klette and Møen (2012) using similar Norwegian data from the years 1982-1995. The point estimates we obtain are somewhat sensitive to specification choices, but there are very few estimates below one. Using log-log specifications our best estimate for the elasticity of R&D with respect to subsidies is about 0.20.

We have demonstrated that there is severe measurement error in the subsidy variable. Additionality is therefore likely to be underestimated. We conclude that measurement errors may be a more important source of bias than selection, and that better subsidy data are imperative in order to improve the additionality estimates.

References

Angrist, Joshua D. and Jörn-Steffen Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arora, Ashish and Alfonso Gambardella (2005). The Impact of NSF Support on Basic Research in Economics, *Annales d'Economie et des Statistiques*, 79-80, 91-117.

Ashenfelter, Orley and Alan Krueger (1994). Estimates of the Economic Return of Schooling from a New Sample of Twins, *American Economic Review*, 84(5), 1157-1173.

Benavente, José Miguel, Gustavo Crespi and Alessandro Maffioli (2007). The Impact of National Research Funds: An Evaluation of the Chilean FONDECYT. OVE Working Paper 03/07, Inter-American Development Bank, Washington, D.C.

Bound, John, Charles Brown and Nancy Mathiowetz (2001). Measurement Error in Survey Data, Ch. 59 in J. Heckman and E. Leamer (eds.) *Handbook of Econometrics*, **5**(5), 3705-3843, Elsevier Science.

Bronzini, Raffaello and Eleonora Iachini (2011). Are incentives for R&D effective? Evidence from a regression discontinuity approach, Bank of Italy Working Papers 791.

Carter, Grace M., John D. Winkler and Andrea K. Biddle-Zehnder (1987). An Evaluation of the NIH Research Career Development Award, RAND Reports,R-3568-NIH.

Cerulli, Giovanni (2010). Modelling and Measuring the Effect of Public Subsidies on Business R&D: A Critical Review of the Econometric Literature, *Economic Record*, 86(274), 421–449.

Chudnovsky, Daniel , Andrés López, Martín A. Rossi and Diego Ubfal (2008). Money for Science? The Impact of Research Grants on Academic Output, *Fiscal Studies*, 29(1), 75-87

David, Paul A., Bronwyn H. Hall and Andrew A. Toole (2000). Is Public R&D a Complement or a Substitute for Private R&D? A Review of the Econometric evidence. *Research Policy*, 29(4-5), 497-529.

Finne, Håkon (2011): Is R&D in the business enterprise sector in Norway under-reported? SINTEF Research Report A20772.

Garcia-Quevedo, José (2004). Do Public Subsidies Complement Business R&D? A Meta-Analysis of the Econometric Evidence, *Kyklos*, 57, 87-102

Griliches, Zvi and Jerry Hausman (1986). Errors in Variables in Panel Data, *Journal of Econometrics*. 31(1), 93-118.

Imbens, Guido W. and Thomas Lemieux (2008). Regression Discontinuity Designs: A guide to practice. *Journal of Econometrics*. 142(2), 615-635.

Jacob, Brian and Lars Lefgren (2011). The Impact of Research Grant Funding on Scientific Productivity. *Journal of Public Economics*. 95(9-10), 1168-1177.

Jaffe, Adam B. (2002). Building Programme Evaluation into the Design of Public Research-Support Programmes, *Oxford Review of Economic Policy*, 18(1), 22-34.

Kauko, Karlo (1996). Effectiveness of R&D Subsidies – A Sceptical Note on the Empirical Literature, *Research Policy*, 25(3), 321-323.

Klette, Tor Jakob and Jarle Møen (2012). R&D Investment Responses to R&D Subsidies: A Theoretical Analysis and a Microeconometric study. *World Review of Science, Technology and Sustainable Development*, 9(2/3/4), 169-203.

Klette, Tor Jakob, Jarle Møen and Zvi Griliches (2000). Do Subsidies to Commercial R&D Reduce Market Failures? Microeconometric Evaluation Studies, *Research Policy* 29(4-5), 471-495.

Lach, Saul (2002). Do R&D subsidies Stimulate or Displace Private R&D? Evidence From Israel, *Journal of Industrial Economics*, 50(4), 369-390.

Lerner, Josh (1999). The Government as Venture Capitalist: The Long Run Impact of the SBIR Program, *Journal of Business*, 72(3), 285-318.

Lichtenberg, Frank R. (1984). The Relationship between Federal Contract R&D and Company R&D, *American Economic Review*, 74(2), 73-78.

Park, Rolla Edward (1966). Estimation with Heteroscedastic Error Terms, *Econometrica,* 34(4), 888.

Serrano-Velarde, Nicolas (2008). Crowding-Out at the Top: The Heterogeneous Impact of R&D Subsidies on Firm Investment. Job market paper, European University Institute.

Wallsten, Scott. 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.

Appendix A: Assessment criteria for the proposals

	Aspect	Evaluated by	Description
A1	General project quality	Expert Panel	General project quality is an expression of how well the project complies with the requirements that should be posed to every project, independent of project content and type. Project quality includes project content and the applicants.
A2	Level of innovation	Expert Panel	The term “innovation” is to be understood in the context of value creation. Evaluation will be focused on the level of innovation compared with the “state-of-the-art” in a field; e.g., at the corporate level, in a particular industry or in a global context.
A3	Research content	Expert Panel	This criterion will be used to rank the project on a scale ranging from simple development work to advanced scientific research. Evaluation will be focused on the extent to which the project produces new knowledge of significance for professional development in the field covered by the research, and the status of the project with regard to the international research frontier.
A4	International cooperation	Research council administrators	Evaluation will be focused on the extent to which the project will contribute to the internationalisation of Norwegian research and/or industry in the relevant field, and the plans for accomplishing this. Furthermore, consideration will be given to whether the selection of international partners will help to enhance the project's quality and feasibility.
A5	Commercial benefits	Expert Panel	Evaluation will be focused on the project's potential benefits for the participating enterprises. The potential refers to anticipated financial gains as a result of industrialisation and commercialisation, and will be compared with the aggregate expenses for the entire period (i.e. beyond the R&D project's duration and expenses per se).
A6	Relevance and benefit to society	Expert Panel	Evaluation will be focused on the extent to which a project is relevant to society; e.g., by considering its ability to contribute to knowledge/competence that would in the short or long term be of significance to meeting major challenges in the public sector, industry and the civil society, viewed in a regional, national or global context.
A7	Risk	Research council administrators	Evaluation will be focused on assessing the extent to which different factors may be expected to inhibit the success of the R&D project, and the planned exploitation of the results.
A8	Other conditions	Research council administrators	Characteristics of the project that are not directly addressed by the other aspects and that could be relevant to the evaluation of the project.
A9	Additionality	Research council administrators	Evaluation will be focused on the extent to which support from the Research Council will trigger inputs, actions, results and effects assumed not to be feasible without the support
A10	Relevance relative to the call for proposals	Research council administrators	The project will be evaluated in relation to the guidelines set out in the call for proposals for the relevant activity/programme.
A11	Total evaluation	Research council administrators	

All aspects except A7 (Risk) and A8 (Other conditions) are evaluated on a scale from 1 to 7.
Source: The Research Council of Norway

Appendix B: Applications of the regression discontinuity design in studies evaluating the effect of R&D support

This appendix briefly summarises five studies evaluating R&D subsidy schemes using a regression discontinuity design. An early regression discontinuity study that pre-dates Jaffe's (2002) article is Carter et al. (1987). Carter et al. evaluate how the careers and accomplishments of Research Career Development Award (RCDA) recipients compare with other researchers. They find that the RCDA programme supports persons who become outstanding researchers, but this is because the programme is able to target them, it is not a causal effect. RCDA is a programme under the US National Institutes of Health (NIH).

Jacob and Lefgren (2011) estimate the impact of receiving grants from the NIH on subsequent publications and citations. Using both OLS and regression discontinuity on a sample of all applications from 1980 to 2000, they find that the receipt of either an NIH postdoctoral fellowship or a standard research grant leads to about one additional publication over the next five years. They suggest that this modest effect on productivity is due to low input additionality. The difference in the number of funding sources between grant winners and losers in their sample is not statistically significant, and hence, it seems that NIH grant receipt displaces other funding. As noted by Jacob and Lefgren, this result does not imply that NIH research funding is not valuable. When overall funding is in fixed supply, support from one specific source may increase the total amount of R&D without having an impact on the marginal applicant.

In a similar study, Benavente, Crespi, and Maffioli (2007) analyse the impact of the Chilean National Science and Technology Research Fund (FONDECYT) on scientific production. Using a regression discontinuity design on projects submitted for funding between 1988 and 1995, they do not find any significant impact of the programme, neither in terms of publications, nor in terms of the quality of publications in the proximity of the programme threshold ranking. Like Jacob and Lefgren (2011) they attribute this lack of success to the possibility that the researchers and projects targeted by the programme have access to alternative funding opportunities.

Serrano-Velarde (2008) uses quantile regressions and regression discontinuity to estimate the impact of R&D subsidies on firm R&D investment under the French ANVAR programme. Rather than using a discontinuity related to proposal evaluation grades, however, Serrano-Velarde utilises a discontinuity resulting from programme specific eligibility requirements related to the form of ownership. He finds that subsidies to large R&D investors crowd out private R&D investment, while subsidies to small R&D investors add to their private investments.

Bronzini and Iachini (2011) also analyse commercial R&D and are in spirit very close to Jaffe (2002) even though he is not cited as a source of inspiration. They evaluate a program for industrial research, innovation and technology transfer put in place in a region of Northern Italy in 2003. Grants are based on an assessment carried out by a committee of independent experts that is appointed by the regional government. Projects that obtain a score of at least 75 on a scale from 0 to 100 receive a grant. This allows Bronzini and Iachini to apply a sharp regression discontinuity design in order to compare the performance of subsidised and non-subsidised firms close to the threshold. They do not have access to data on the firms' R&D investments, but use balance sheet variables that are associated with R&D outlays reimbursable by the program. More specifically, they use intangible, tangible and total investments as their main outcome variable, and also look at labour costs, employment level and service costs. By letting the outcome variables be a function of the score, the average treatment effect of the program is assessed through the estimated value of the discontinuity at the threshold. They use both parametric and non-parametric methods. When using the full sample of firms, they cannot reject the hypothesis that firms substitute public for privately financed R&D, in which case the program does not create additional investments. However, when they explore heterogeneity, they find evidence suggesting that small firms do increase their investments substantially, on average by the same amount as the subsidy they receive.

Appendix C: Accounting for R&D intentions using an instrumental variables approach

In addition to the two approaches reported in the main article, we have also tried to solve the selection problem using instrumental variable estimation. This approach did not succeed, but we report the results here for completeness. The idea is as follows. Conditional on commercial benefits, evaluation Aspect 6, 'Relevance and benefit to society', should not affect the private decision to invest in R&D. It should, however, affect the decision of the granting authority, both in terms of whether or not to grant a subsidy, and in terms of the amount granted. We also try using the total evaluation grade (MG11) as an instrument because it captures other aspects of the proposal that are valued by the granting authority. We stress that this assumption contradicts the assumption motivating the inclusion of MG11 in Table 8, that MG11 might capture something more than MG5 about the value of projects for firms.

Table C1. IV estimation for endogenous subsidy. Levels form

	Pooled OLS			Within (FE) estimator		
Instrument:	none	MG6	MG11	None	MG6	MG11
		(1)	(2)		(4)	(5)
S^R (endogenous)	2.756**	14.080*	8.434	1.275***	-1.228	.586
	(1.181)	(7.404)	(9.930)	(.433)	(3.371)	(1.613)
S^{EU}	7.947***	2.660	5.298	2.727	3.178	2.867*
	(1.780)	(4.275)	(5.004)	(1.692)	(1.975)	(1.737)
S^G	1.486***	1.454***	1.470***	.340***	.349***	.343***
	(.028)	(.0628)	(.0557)	(.0753)	(.0825)	(.0786)
Sales	3.946***	3.796***	3.870***	.539	.539	.532
	(.920)	(.846)	(.890)	(.704)	(.702)	(.703)
Sales squared	-.00628*	-.00647*	-.00638*	.00240	.00253	.00245
	(.00363)	(.00333)	(.00346)	(.00212)	(.00214)	(.00211)
MG5	2262***	283.5	1270	192.8	554.5	325
	(440)	(1110)	(1856)	(179.7)	(470.4)	(287.8)
MG5MISS	3537*	-2115	704.2	293.5	638.4	532.3
	(1812)	(4189)	(5146)	(873.3)	(954.7)	(1009)
REJECT	-2125	5620	1757	-701.2	-1820	-1078
	(1817)	(5048)	(7294)	(1165)	(1665)	(1560)
Coefficients on instruments in first stage:						
MG6		184.6***			152.7**	
		(62.10)			(71.13)	
MG11			43.75*			-27.18
			(26.34)			(53.93)
N	11368	11368	11368	11368	11368	11368
Number of firms	2570	2570	2570	2570	2570	2570
R-squared	.368	.265	.342	.037	.019	.036

The dependent variable is intramural R&D. All specifications include year dummies. Pooled OLS also includes dummies for 2-digit NACE group. Weighted using Park's optimal weight. Standard errors allowing for clustering of residuals by firm are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C2. IV estimation for endogenous subsidy. Log-log form

	Pooled OLS-IV			Fixed effects-IV		
Instrument:	none	MG6	MG11	none	MG6	MG11
	(1)	(2)	(3)	(4)	(5)	(6)
ln S^R (endogenous)	.306***	.665**	.246	.220***	.0338	-.0851
	(.0217)	(.288)	(.491)	(.0215)	(.448)	(.650)
ln S^{EU}	.135***	.00868	.156	.0675**	.111	.146
	(.0294)	(.107)	(.176)	(.0294)	(.139)	(.200)
ln S^G	.348***	.316***	.353***	.305***	.303***	.311***
	(.0158)	(.0312)	(.0464)	(.0220)	(.0396)	(.0529)
ln sales	.329***	.319***	.330***	.301***	.268***	.270***
	(.0359)	(.0371)	(.0380)	(.0712)	(.0714)	(.0729)
MG5	.289***	.0632	.326	.00344	.0677	.119
	(.0418)	(.185)	(.311)	(.0431)	(.202)	(.290)
MG5MISS	1.098***	.284	1.234	-.423	-.256	-.101
	(.274)	(.697)	(1.147)	(.285)	(.651)	(.899)
REJECT	.0741	.951	-.0729	.271	-.0202	-.227
	(.215)	(.729)	(1.220)	(.205)	(.806)	(1.154)
Coefficients on instruments in first stage:						
MG6		.422***			.241**	
		(.101)			(.115)	
MG11			.164**			.122*
			(.0639)			(.0684)
N	11368	11368	11368	11368	11368	11368
Number of firms	2570	2570	2570	2570	2570	2570
R-squared	.221	.207	.221	.059	.066	.057

The dependent variable is intramural R&D. All specifications include year dummies. Pooled OLS also includes dummies for 2-digit NACE group. Standard errors allowing for clustering of residuals by firm are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We apply the IV models both to the pooled OLS and to the fixed effects models, while maintaining the proxy variable approach. Table C1 shows the results for the levels equations while Table C2 shows the results for the log-log form. The coefficients of the instruments in the first stage regressions are at the bottom of the tables. For comparison we include the results from the corresponding regressions in Table 8 and 10 where subsidies were not instrumented. Generally, the instruments do not work well. The instruments are significant in the first stage regressions in seven out of eight specifications, but the estimated additionality coefficients tend to be either implausible high (OLS) or insignificant in the second stage regressions (FE). We have also tried combining the two instruments without having any more success.

B

Return to:
Statistisk sentralbyrå
NO-2225 Kongsvinger

From:
Statistics Norway

Postal address:
PO Box 8131 Dept
NO-0033 Oslo

Office address:
Kongens gate 6, Oslo
Oterveien 23, Kongsvinger

E-mail: ssb@ssb.no
Internet: www.ssb.no
Telephone: + 47 62 88 50 00

ISSN 0809-733X

Statistisk sentralbyrå
Statistics Norway

Design: Siri Boquist