Are Complementary Policies Substitutes? Evidence from R&D Subsidies in the UK

Jacquelyn Pless*

September 30, 2023

Abstract

This paper studies whether grants and tax incentives for private R&D are complements or substitutes. I use multiple quasi-experimental research designs to examine firms in the United Kingdom and find that increasing tax credit generosity substantially enhances the effect of grant funding on R&D for small firms, suggesting that the instruments are complements. Financial constraints are likely at play. The effects are strongest for firms that appear constrained, and the combination of policies increases R&D "entry." Furthermore, I find that the instruments are substitutes for larger firms, which are usually less constrained. Some alternative explanations can be ruled out.

Keywords: R&D; innovation; policy interactions; difference-in-discontinuities

JEL codes: D22, H0, H25, L53, O31, O32, O38

*Assistant Professor at MIT Sloan School of Management, 100 Main Street, Cambridge, MA 02142. Phone: (617) 715-4849. Email: jpless@mit.edu. I wish to thank Pierre Azoulay, Nick Bloom, Francois Cohen, Antoine Dechezleprêtre, Simon Dietz, Niall Farrell, Naomi Feldman, Harrison Fell, Meredith Fowlie, Ben Gilbert, Irem Guceri, Cameron Hepburn, Sabrina Howell, Dan Kaffine, Ian Lange, Ashley Langer, Josh Lerner, Danielle Li, Pete Maniloff, Linus Mattauch, Myra Mohnen, Pierre Mohnen, Will Rafey, Mark Smout, Sugandha Srivastav, Scott Stern, Rich Sweeney, Chad Syverson, Alex Teytelboym, Arthur van Benthem, and nine anonymous referees for helpful feedback. I also thank participants in seminars at MIT Sloan School of Management, University of California-Berkeley, Boston University, University of Wisconsin-Madison, Imperial College London, University of Maryland, Ivey Business School at Western University, Brookings Institution, Georgia Tech, University of Oxford, Harvard University, University of Sussex, Colorado School of Mines, University of Kansas, NBER Summer Institute, Roundtable for Engineering Entrepreneurship Research, IIPF Annual Conference, CESifo Area Conference on Energy & Climate Economics, and AERE Summer Conferences for comments. The UK Data Service Secure Lab helped make this research possible by providing access to data. I gratefully acknowledge financial support from the Oxford Martin Programme on Integrating Renewable Energy and Partners for a New Economy. I remain fully responsible for the paper's contents and any mistakes herein.

1 Introduction

Innovation is a central driver of economic growth and competitiveness, but it tends to be under-supplied by markets due to knowledge spillovers (Nelson 1959; Arrow 1962). This market failure underpins the economic rationale for intervention, and indeed, governments globally spend hundreds of billions of dollars on subsidies for research and development (R&D) every year. Yet designing innovation incentives so they live up to their promise remains a long-standing and increasingly pressing challenge (Bloom, Van Reenen and Williams 2019). As policymakers worldwide are revisiting their innovation strategies in an effort to revitalize economies following the COVID-19 pandemic while also tackling ongoing widespread crises like climate change, developing a better understanding of how to foster innovation is increasingly urgent. With many countries facing productivity growth declines since the 1970s, industrial policy is also back in fashion, and financial support for R&D is at the forefront.[1]

Governments typically use a mix of programs and policies to subsidize private R&D. Direct grants and tax incentives are particularly popular, and there is growing evidence that they both boost innovative activity. However, relatively little is known about their effects in the context of the broader innovation policy ecosystem, which is important to consider when designing incentives. That is, do instruments like grants and tax credits interact in their effects on firm behavior? Firms frequently tap into multiple support schemes, and as the choices organizations make in response to one incentive might depend on the availability and generosity of others, policy interactions could either enhance or dampen the marginal return to each instrument depending on whether they are complements or substitutes.

Consider a firm lacking resources to finance the start-up costs associated with developing a new technology, like purchasing expensive machinery or setting up a lab. An upfront grant could help the firm overcome this barrier while also freeing up resources for R&D labor (e.g., scientists and engineers). Tax credits often can be used to subsidize labor-related expenditures, so if firms hire more researchers when tax credits become more generous, grant-funded capital may become more productive. This would imply that the two instruments are

[1]President Biden's 2024 Budget included $210 billion for Federal R&D, an all-time high, including $101 billion for basic and applied research (OMB 2023). The US Inflation Reduction Act (IRA) also directs more than $300 billion specifically towards clean energy investments. The UK's Industrial Strategy also targets specific industries to tackle four "grand challenges."

complements. On the other hand, for firms that can self-finance upfront costs of a project, increasing the generosity of tax credits may make grants less attractive. The benefits of tax credits may outweigh the value of grants (minus the cost and burden of applying). Less constrained firms may choose to no longer apply for grants, diminishing the contribution of grant-funded projects to the total R&D. This would suggest the subsidies are substitutes.

In this paper, I implement six quasi-experimental research designs to present new evidence on whether R&D grants and tax credits are complements or substitutes in their effects on firms' R&D expenditures. Studying this question is empirically challenging because it requires variation in both support schemes and randomization in innovation policy is rare. I overcome this by exploiting several sources of policy-induced variation in the cost of investing in R&D for firms in the United Kingdom. I primarily focus on small firms, which can play a pivotal role in driving innovation but are often constrained, and they tend to be hard to study due to data limitations. I also examine larger firms when exploring mechanisms.

I primarily study firms that receive grants from Innovate UK, the UK's premier public funding agency for private sector innovation, which spent £885 million in fiscal year 2020/2021. To start, I estimate the independent effect of grant funding on small firms' R&D expenditures using a discontinuity in "funding rates" (i.e., the proportion of project costs that is subsidized). Innovate UK applies different pre-determined funding rates based on firm size, whereby firms under specific employment, total assets, and turnover thresholds benefit from funding rates that are, on average, 7.5 percentage points (or 15.5%) higher relative to firms over the size thresholds. Then, to study how grant funding generosity interacts with tax credits, I use a difference-in-discontinuities approach (henceforth "diff-in-disc") that essentially entails testing whether there is a change in the discontinuity at the grant funding threshold when tax credit rates under the UK's R&D Tax Relief for Corporation Tax Scheme (henceforth "tax credit") increase over time.

There are a few key conditions for these research designs to produce unbiased estimates. The first set includes the standard RDD assumptions of threshold independence, no manipulative sorting around the cutoff, and continuity in potential outcomes. The diff-in-disc assumptions are similar but require there to be no *changes* in these factors when tax credit rates increase. I probe these assumptions with numerous empirical tests and also manually review other policies in the UK and the assumptions appear to hold.

The headline result is that direct grants and tax credits are complements in their effects on R&D for small firms. While the 15.5% difference in grant funding rates alone already doubles R&D expenditures for firms just below the threshold relative to those over it when examining the independent effect across the full sample period (2005-2012), the effect increases substantially when moving from a "low" tax credit rate period (2005-2012) to a "high" tax credit rate period (2013-2017). In my preferred specification, I find that the 40% increase in tax credit benefits enhances the discontinuity in R&D around the grant funding rate threshold by about £560k, which is 2.8 times the size of the independent effect estimate. The results are robust to a variety of modeling assumptions and when increasing the bandwidth to include a much wider range of firm sizes, suggesting that the phenomenon may not be limited to just a narrow set of firms around the threshold.

What can explain such large effects? Evidence from further analyses suggest that the combination of subsidies helps firms overcome financial constraints, which may contribute to the substantial increases in R&D. For example, if the available of both subsidies enables firms pursue new projects that require purchasing new machinery or setting up a lab, R&D expenditures will increase substantially due to the high upfront costs. I provide three sets of results that are consistent with financial constraints being at play. First, when splitting the sample based on variables that proxy for financial constraints, I find that the effects are much larger for firms that appear constrained. Second, the policy interaction induces R&D "entry" and increases the propensity to invest for firms that are already R&D performers. On the other hand, more generous grant funding on its own does not have these extensive margin effects. This suggests that, while increasing the generosity of grants alone has an effect on R&D on the intensive margin, it also may increase the number of R&D firms in the economy but only in the presence of higher tax credit rates.[2]

Third, to corroborate the conclusion that financial constraints are driving subsidy complementarity for small firms, I estimate the effects on larger firms' R&D, as they are less frequently constrained. One might expect no complementarity if financial constraints are

[2]This is consistent with Mohnen and Röller (2005)'s conclusion that a *package* of innovation policies can increase the propensity to innovate.

indeed at play. I compile additional administrative data and start by estimating the independent effect of tax credits to establish a baseline that is comparable to other studies in the literature examining the same policy. I use a discontinuity in tax credit rate generosity, whereby firms with fewer than 500 employees are eligible for much higher rates relative to firms above the cutoff, and implement an RDD using post-policy (2009-2014) data and a diff-in-disc design both pre- and post-policy data (2000-2014). The results from both approaches indicate that being eligible for more generous tax credits has a large, positive effect on R&D expenditures, consistent with existing studies of the tax credit policy on its own (Guceri and Liu 2019; Dechezleprêtre, Einiö, Martin, Nguyen and Van Reenen forthcoming).

Next, to examine policy interactions for larger firms, I use research designs that I refer to as "discontinuity-in-effects" and "difference-in-discontinuity-in-effects" approaches that test whether more generous tax credit rates dampen or enhance the correlation between direct subsidy funding and R&D. With a difference in the correlation between subsidy funding and R&D being driven strictly by the exogenous tax credit generosity threshold for firms within a narrow window around the cutoff, I interpret a dampening of the correlation for firms under the threshold as a substitution effect, and conversely, an enhancement as complementarity. I also extend these methods by incorporating pre- and post-policy variation. The results from all four approaches indicate that the correlation between direct subsidy funding and R&D is substantially lower for firms just below the tax credit generosity threshold relative to those just over it, suggesting that the instruments are substitutes for larger firms.

I explore alternative explanations of the positive interaction effects for small firms as well but do not find evidence that supports them. For example, there is no shift in the type of research that is funded (which could inflate the estimates if grant funding is more effective for specific types of projects) and there does not appear to be preferential treatment of previous grant winners, which could enhance the marginal effect of grant funding over time due to cumulative funding. Substitution away from grants when tax credit rates increase also does not seem to drive the results based on indirect tests of changes in competition, but conclusions on this point should remain cautious without having data on all applicants.

This paper contributes to the literature estimating the causal impacts of innovation and industrial policies in a few ways. First, it connects two separate but related literatures using

RDDs and other quasi-experimental methods to examine the effects of direct grants and tax credits. For example, Bloom, Griffith and Van Reenen (2002), Rao (2016), Guceri and Liu (2019), Dechezleprêtre et al. (forthcoming), and Agrawal, Rosell and Simcoe (2020) find that tax incentives increase R&D, and Bronzini and Iachini (2014), Howell (2017), and Azoulay, Graff Zivin, Li and Sampat (2018) find that grants have positive effects on patenting. While existing work focuses on the independent effect of each instrument, I extend the scope of the evidence base by studying the interactions of these instruments, which impact the returns to subsidies and elucidate the importance of removing silos when designing policy.

Relatedly, my findings suggest that using a *combination* of policies may be important for inducing non-R&D performers to "enter" and to increase the propensity of firms to invest on the extensive margin. This is not to say that receiving a grant does not induce R&D entry. Rather, increasing the generosity of grant rates (and thus levels) might also have this effect, but only in the presence of higher tax credit rates as well. In carrying out these analyses, I provide new results to the literature specifically on how innovation subsidies impact small firms, which contribute disproportionately to major innovations and grow faster than larger firms (Akcigit and Kerr 2018) but are often constrained and can be difficult to study due to data limitations.[3]

More broadly, my findings and methods also relate to the literature studying the effects of business support programs and policies, especially in light of how firm size-based policies are common in many economic settings. Lastly, policy interactions are ubiquitous but there is limited, well-identified evidence of their effects. This paper therefore also may be of interest to other fields for which policy interactions are prevalent.[4]

The remainder of this paper is organized as follows. Section 2 provides institutional details. Sections 3 and Section 4 describe the research designs and data, and Section 5 probes the identification assumptions. The main results are presented in Section 6. Section 7 explores the underlying mechanisms and I conclude in Section 8.

[3]One notable exception is Agrawal et al. (2020)'s study of tax credits in Canada. Note that while Guceri and Liu (2019) and Dechezleprêtre et al. (forthcoming) study the UK's tax credit scheme for SMEs, SMEs are defined as including firms up to 500 employees for this policy whereas the sample I use through most of my small firm analyses includes firms with fewer than 100 employees.

[4]There is a literature examining whether information interventions and market-based tools are complementary (Duflo, Dupas and Kremer 2012; Ashraf, Jack and Kamenica 2013; Dupas 2009), and on the complementarity of programs impacting labor supply (Inderbitzin, Staubli and Zweimuller 2016).

2 Institutional Details

2.1 Direct Grant Funding from Innovate UK

Innovate UK, a non-departmental public body, has been the UK's premier grant-awarding agency for business-led innovation since 2004. It receives a fixed budget allocation from the Department of Business, Energy and Industrial Strategy (BEIS) each year and has provided more than £15 billion in direct funding to the private sector since its inception (InnovateUK 2023). Grants are awarded through competitions managed by Innovate UK, which are often sector-specific of mission-driven but they are also sometimes more open, calling for any novel innovations that have potential to make a "significant impact on the UK economy." Most commonly, a specific budget for each competition is set upfront, and when firms apply for funding, they include proposed project costs. If firms are successful, Innovate UK subsidizes a proportion of total project costs following the rules detailed below.

The typical application and evaluation process is as follows.[5] Applicants submit proposals detailing the scope of the project, including costs, timelines, and planned activities. Applications are screened for meeting the competition's criteria and allocated to three to five independent assessors from the private sector and academia based on skill sets and areas of expertise. Assessors submit scores based on standardized questions and provide feedback, which Innovate UK compiles to identify a ranked order of all applications based on the average scores. An Innovation UK staff member responsible for the competition moderates and reviews the rankings. There is often an interview stage once a shortlist is identified when scores can be updated and re-ranked. The Innovate Lead for the competition then typically recommends the highest-ranked applications for funding, but sometimes they may recommend a portfolio of projects to align with the competition's (if stated in the competition description). I discuss this more in Section 7.2.1.

The main feature of the program that I employ is a discontinuity in grant funding "rates" (i.e., the proportion of proposed project costs that is subsidized). Innovate UK guidelines created early during the program's launch set different grant funding rates based on firm size, whereby firms under specific employment, turnover, and total assets thresholds that are

[5] A more detailed account of the process is detailed in UKRI (2023).

used by the European Commission for defining whether firms are small, medium, and large benefit from more generous rates than those over the thresholds. I focus on the "small" firm thresholds, whereby firms are defined as small if they have fewer than 50 employees and either a maximum turnover or balance sheet total of €10m. Small firms are eligible for 70 percent, 70 percent, and 45 percent of total project costs to be subsidized for feasibility studies, industrial research, and experimental development projects, respectively. On the other hand, those just over the threshold are eligible for funding that subsidizes 60 percent, 60 percent, and 35 percent of project costs, respectively. Fundamental research projects are 100 percent funded for firms of all sizes, so the average difference in grant funding rates for firms just under the small firm size threshold is 15%, or 7.5 percentage points.

2.2 R&D Tax Relief Scheme

The UK's R&D Tax Relief for Corporation Tax Scheme (henceforth "R&D tax credit"), introduced in 2000 for small- and medium-sized enterprises (SMEs) and extended to large companies in 2002, is volume-based, applying enhanced deductions of current R&D expenditures from taxable income, thus reducing corporate tax liabilities.[6] It is a particularly generous incentive in the UK and used widely, as applying is relatively straightforward, entailing the completion of a supplementary form when filing company tax returns. Loss-making firms can also benefit through a payable tax credit. Since its launch, more than £16.5 billion in tax relief has been claimed under the R&D tax credit scheme, with £2.9 billion spent in fiscal year 2015/16 alone, (HMRC 2017), accounting for more than 80% of government support for business R&D in 2019.

I use two key sources of policy-induced variation in my analyses. First, when studying policy interactions for small firms, I use variation in tax credit rates over time, distinguishing between a "low tax credit rate period" (2005 through 2012) and a "high tax credit rate period" (2013 through 2017). Second, when studying larger firms, I exploit a discontinuity in tax credit rates, whereby firms qualifying as SMEs under the policy benefit from much

[6] This is in contrast to incremental R&D tax incentives, such as those in the U.S, where firms benefit only if their R&D expenditures exceed a base level of previous expenditures. The main benefit that the volume-based design offers is simplicity, and thus it is widely used by firms of all sizes and ages in the UK.

more generous tax credits. Starting in 2008, the R&D tax credit policy defined firms as SMEs if they have fewer than 500 employees and either no more than €100m in sales or no more than €86m in total assets. Importantly, these thresholds apply only for the tax credit policy—for all other intents and purposes, firms must have fewer than 250 employees (and meet other turnover and total assets criteria) to qualify as SMEs, as was the case for the R&D tax credit policy prior to 2008. As such, while firms up to 499 employees are defined as SMEs under the R&D tax relief program, I refer to them as "larger."[7]

Table 1 details the components determining tax credit benefits from 2005 through 2017 (enhancement and corporate tax rates) along with the percentage of R&D expenditures that the tax credit subsidizes based on the policy's formula.[8] Between 2005 and 2017, the enhancement rate increased from 0.75 to 1.30. I exploit changes over time in my empirical analyses. The first major increase happened in 2011, followed by another big increase in 2012 and smaller changes in later years. I split the years into a pre-tax credit change period (2005-2012) and post-tax credit rate change period (2013-2017) for the empirical analyses and use the average tax credit benefit in the two periods (i.e., the percentage of R&D expenditures that is subsidized according to the formula described above).[9] The average benefit increased by 33% from 16.5% of expenditures in the pre-tax credit change period to 22% afterwards.

Appendix Table C.1 provides the enhancement rates for firms that are over the R&D tax credit SME thresholds from 2000 to 2014. On average, the proportion of R&D that is subsidized for firms over the thresholds is 17 percentage points lower than it is for those under them.

[7]Dechezleprêtre et al. (forthcoming) use these thresholds to study the effects of the tax credit policy on its own, showing that the more generous rates have positive effects on R&D expenditures and patenting. I use them when examining the tax credit interaction with grant funding.

[8]The percentage of R&D expenditures that the tax credit subsidizes is equal to the product of the enhancement rate and the corporate tax rate. The same formula applies for loss-making firms except that the corporate tax rate is replaced with the payable credit rate.

[9]I take 2013 as the first post-change year to account for how a policy change in 2012 would affect the incentive to invest in fiscal year 2012-13, and the 2011 and 2012 changes were both large yet followed by only incremental changes.

3 Empirical Strategies

I employ six quasi-experimental empirical strategies to provide evidence on the effects of innovation grant funding, R&D tax relief, and the interactions of the two policy instruments for firms in the UK. I focus primarily on small firms, first using a regression discontinuity design (RDD) to estimate the independent marginal effect of higher grant funding rates over the entire sample period and then a difference-in-discontinuities ("diff-in-disc") approach to study the effects of grant and tax credit interactions. When exploring the underlying mechanisms behind the results for small firms, I also examine larger firms by implementing additional regression discontinuity and diff-in-disc designs to estimate the independent effect of being eligible for more generous tax credits and then two extensions to these methods—"discontinuity-in-effects" and "difference-in-discontinuities-in-effects—approaches to study policy interactions for larger firms.

To produce unbiased estimates, each of the research designs relies on different but closely related identifying assumptions. I discuss and probe these assumptions in Section 5.

3.1 Research Designs to Study Small Firms

3.1.1 RDD Approach to Estimate the Effect of Grant Funding

I start by estimating the independent marginal effect of grant funding on Innovate UK grant winners' R&D expenditures over the entire sample period (2005-2017). This helps establish a baseline before moving to the policy interactions. To do so, I employ a regression discontinuity design (RDD) approach using the 50-employee cutoff determining whether firms are eligible for more generous grant funding rates (i.e., a higher proportion of their proposed project costs are subsidized) relative to firms just over the threshold. The running variable is centered employment in the year prior to receiving a grant. The main identifying assumptions are that that firms receiving Innovate UK grants just above the cutoff provide a good comparison group for those just below the cutoff and that firms do not precisely manipulate their size around the threshold.

More formally, centered employment (A^*_{it}) for firm i in year t is the firm's employment in the year prior to receiving a grant minus 50 (i.e., $A^*_{it} = A_{it} - A_c$). Firms are defined as treated if the running variable is below 50 employees and the other turnover and total assets eligibility criteria described in Section **??** are also met. I estimate a local linear regression of the following form:

$$Y_{it} = \delta_0 + \delta_1 A^*_{it} + J_{it}(\gamma_0 + \gamma_1 A^*_{it}) + \eta_{st} + \mathbf{X}_{it}\phi + \varepsilon_i, \tag{1}$$

where Y_{it} is the outcome of interest for firm i in year t (which is primarily R&D expenditures throughout this paper), J_{it} is an indicator equal to one if the firm is treated, and ε_{it} is the random error. The slope of the running variable is allowed to differ on either side of the cutoff.[10] Furthermore, while the RDD does not require additional controls, including them can improve precision and help ensure that estimated coefficients are not contaminated by pre-existing differences between treated and untreated firms. I therefore also estimate variations of the model that sometimes include industry-year fixed effects (η_{st}) and different sets of time-varying firm-level controls ($\mathbf{X}_{it}$).

The main parameter of interest is γ_0, which identifies the local average treatment effect of being eligible for more generous grant funding rates on the outcome of interest for small firms under a specific set of assumptions detailed in Section 5. Throughout the main analyses, the estimating sample includes observations for the year in which the firm receives a grant as well as the three years that follow. As such, I apply the treatment definition not only to the year in which firms receive grants but also the three that follow for each grant that it receives (as opposed to updating the running variable and treatment status each year, as it is the firm's size when applying for a grant that determines the funding rates).[11] If the firm receives another grant within those three years, the lagged values and treatment status associated with the new grant then replace the originals. I apply a local linear regression to observations within bandwidth h on each side of the employment cutoff A_c, restricting the sample to firms within the interval $[A_c - h, A_c + h]$ using the MSE-optimal bandwidth.[12]

Although eligibility for higher grant funding rates is determined not only by employment but also turnover and total assets thresholds, I use employment as the running variable

[10] In other words, the running variable (A^*_{it}) is included on its own and then also interacted with J_{it}.

[11] I do the same for the running variable controls as well.

[12] I describe the bandwidth selection at the end of Section 4.1.

because it is the binding criteria (i.e., firms must have fewer than 50 employees, and if so, they must also have no more than €10 million in turnover or total assets). I account for the other criteria by limiting the estimation sample to include only those that also fulfill the turnover or total assets eligibility requirements for more generous grant funding rates.[13]

3.1.2 Difference-in-Discontinuity Approach to Examine Policy Interactions

To test whether grant funding and tax credits are complements or substitutes for small firms, I exploit increases in R&D tax credit rates over time and take a difference-in-discontinuities (henceforth "diff-in-disc") approach that involves incorporating elements of difference-in-differences methods to the previous RDD model. The intuition behind the research design essentially is that it tests whether the discontinuity in R&D at the grant funding rate threshold changes when tax credit rates increase. More specifically, I estimate the difference in the discontinuity when shifting from a "low tax credit rate period" (2005-2012) to a "high tax credit rate period" (2013-2017) using variations of the following (local) linear model:[14]

$$\begin{aligned} Y_{it} = \delta_0 + \delta_1 A^*_{it} + J_{it}(\gamma_0 + \gamma_1 A^*_{it}) + T_t[\alpha_0 + \alpha_1 A^*_{it} + J_{it}(\beta_0 + \beta_1 A^*_{it})] \\ + \mathbf{X}_{it}\phi + \omega_i + \eta_{st} + \varepsilon_{it}, \end{aligned} \tag{2}$$

where Y_{it}, J_{it}, A^*_{it}, η_{st}, $\mathbf{X}_{it}$, and ε_{it} are defined as before and the main extension relative to the RDD is the introduction of T_t, an indicator equal to one in the high tax credit period and zero otherwise. I also include firm fixed effects (ω_i) in some specifications to control for unobservable time-invariant differences across firms. The baseline sample again includes firms only within the MSE-optimal bandwidth around the 50-employee cutoff (that also meet either the turnover or total assets grant rate generosity criteria) for years in which firms receive a grant and the three years that follow.

The main coefficient of interest is β_0, the diff-in-disc parameter that captures the subsidy interaction effect under the set of assumptions that I discsus below. If β_0 is positive, this indicates that increasing tax credit generosity enhances the effect of grant funding, suggesting

[13] I also show later that the results are robust to not imposing this restriction.

[14] See Grembi, Nannicini and Troiano (2016) for the formalization of the diff-in-disc estimator for identifying the (local) average treatment effect of interactions like this.

that the two subsidies are complements. If β_0 is negative, higher tax credit rates dampen the marginal effect of grant funding, suggesting they are substitutes. Finding no interaction effect would imply, of course, that they do not interact in their effects on firm R&D.

3.2 Research Designs to Study Larger Firms

The key source of variation that I rely upon to study larger firms in all four empirical approaches is the discontinuity in R&D tax credit rates at the SME employment threshold, whereby the tax credit policy's thresholds for defining SME status are double those used for all other intents and purposes in the UK. That is, firms with fewer than 500 employees are eligible for more generous R&D tax relief (conditional on also meeting additional turnover and total assets criteria), generating an exogenous difference in the cost of investing in R&D at the cutoff.[15] I start by using the threshold to estimate the independent tax policy effect and then extend the models to study interactions with direct subsidies.

3.2.1 Two Approaches to Study Independent Effect of Tax Credits

Regression Discontinuity Design. My first approach to study the independent effects of tax credits is an RDD. I estimate models analogous to Equation 1 but replace J_{it} with another indicator variable, C_{it}, equal to one when firms are under the 500-employee tax credit generosity threshold and zero otherwise. I restrict the sample to a narrow window around the threshold but do not limit it to only include firms that also meet the other criteria due to data limitations.

Difference-in-Discontinuity Design. I also implement another diff-in-disc research design that introduces time variation based on when the tax credit firm size thresholds doubled (as described in Section 2). In this case, I estimate how the discontinuity in R&D at the 500-employee tax credit generosity threshold changes when going from the pre-policy period (2000-2008) when there should be no such discontinuity to the post-policy period (2009-

[15] Dechezleprêtre et al. (forthcoming) and Guceri and Liu (2019) also study this policy to examine the effects of tax credits on their own. They refer to these firms as SMEs, since they are defined as such for the R&D tax credit purposes. I refer to them as larger firms here because they are large firms by all other definitions in the UK and EU and much larger than those in the small firm sample.

2014). The model is analogous to Equation 2 but I again replace J_{it} with C_{it} and now also replace T_t with TC_t, an indicator equal to one in the post-policy period and zero otherwise.

3.2.2 Two Approaches to Study Policy Interactions for Larger Firms

To examine policy interactions for larger firms, I extend the previous two methods by interacting the tax credit treatment variables with the amount of funding the firm receives in direct subsidies for R&D, S_{it}. The idea is to test whether there is a discontinuity in the correlation between subsidy funding and R&D at the tax credit generosity threshold. I refer to these approaches as "discontinuity-in-effects" and "difference-in-discontinuities-in-effects."

Discontinuity-in-Effects. Building from the baseline RDD and still limiting the sample to include only firms within a narrow window around the cutoff, I interact the indicator for whether the firm is eligible for more generous tax credits (C_{it}) with the amount of direct subsidy funding for R&D (£000s) that the firm received (S_{it}):

$$Y_{it} = \delta_0 + \delta_1 A^*_{it} + C_{it}(\gamma_0 + \gamma_1 A^*_{it}) + \beta_1(C_{it} \times S_{it}) + \beta_2 S_{it} + \mathbf{X}_{it}\phi + \varepsilon_{it}, \tag{3}$$

where Y_{it}, C_{it}, A^*_{it}, $\mathbf{X}_{it}$, and ε_{it} are defined as before. The primary coefficient of interest is β_1, which captures the discontinuity in the correlation between grant funding and R&D (i.e., the policy interaction effect). If β_1 is positive, the correlation between subsidy funding and R&D is enhanced for firms eligible for more generous R&D tax credits just below the threshold relative to those just over the threshold. This would suggest that the instruments are complements. On the other hand, if β_1 is negative, being eligible for more generous R&D tax credits dampens the correlation between direct subsidy funding and R&D.

The main identifying assumptions behind this research design are similar to those associated with the RDD (i.e., continuity in potential outcomes, no manipulative sorting, and no confounding policies) but with one addition for the interaction component. Note that, while the causal effect of direct subsidy funding cannot be identified without a valid instrumental variable, any existing *discontinuity* in the correlation at the threshold should be driven only by the difference in tax credit rates. As such, as long as the endogeneity of direct subsidy funding goes in the same direction and has a similar magnitude for firms just below and above the tax credit threshold, the difference in the correlation between direct subsidy funding and R&D (i.e., the interaction) at the tax credit threshold should identify the tax credit interaction. I probe this and the other identifying assumptions in more detail in Section 7.

Difference-in-Discontinuities-in-Effects. Finally, I combine this discontinuity-in-effects model with elements of the diff-in-disc approach by incorporating before/after tax credit policy implementation time variation. This entails interacting the tax credit rate threshold indicator and subsidy funding variables with TC_{it}:

$$Y_{it} = \beta_1(C_{it} \times S_{it} \times TC_{it}) + \beta_2(C_{it} \times TC_{it}) + \beta_3(C_{it} \times S_{it}) + \beta_4(S_{it} \times TC_{it}) + \beta_5 S_{it}$$
$$+\delta_0 + \delta_1 A^*_{it} + C_{it}(\gamma_0 + \gamma_1 A^*_{it}) + TC_t[\alpha_0 + \alpha_1 A^*_{it} + C_{it}(\alpha_2 + \alpha_3 A^*_{it})] + \mathbf{X}_{it}\phi + \gamma_i + \eta_t + \varepsilon_{it}, \tag{4}$$

where all variables are as defined before.[16] The main coefficient of interest is β_1, capturing the difference in the discontinuity in the correlation between direct subsidy funding and R&D expenditures in the post-policy period (2009-2014) relative to the pre-policy period (2000-2008). Another coefficient of interest is β_2, the diff-in-disc estimate for the independent effect of the tax credit policy. The interaction between C_{it} and S_{it} accounts for any discontinuities in the correlation in the pre-policy period and the interaction between S_{it} and TC_{it} accounts for how the correlation between direct subsidy funding and R&D may have been different in the pre- and post-policy time periods for other reasons beyond the tax credit policy change.

4 Data and Sample Construction

4.1 Small Firms

To study small firms, I start with grant-level data from Innovate UK's public Transparency Database containing information such as application and funding dates, competition reference numbers, award amounts, and proposed project costs. It also includes company registration numbers (CRNs) so that firms can be uniquely identified and matched over time

[16] Some of the double interactions and main effects are absorbed when including the full set of fixed effects. The running variable components are also all interacted with S_{it} but are omitted here for conciseness.

as well as to other firm-level data sets. I link the Innovate UK data to Bureau van Dijk's Financial Analysis Made Easy (FAME) database for detailed information from firm balance sheets and profit and losses (P&L) statements, such as R&D expenditures and employment, as well as other financial performance variables and time-invariant firm characteristics.[17] Appendix A provides details on the steps taken to prepare each data set before merging.

Of the 12,072 observations from 2005 through 2017 in the prepared Innovate UK grant data set (firm-year observations for the years in which firms receive grants), 11,750 observations (97%) across 7,146 organizations match to FAME on company registration numbers. I apply a few standard rules to check for errors in the financial data, such as omitting observations for which key variables are negative when they should not be, providing an unbalanced panel of 72,481 observations across 7,035 firms from 2005 through 2017.

In my baseline estimations, I restrict the data to include just the year in which firms receive a grant and the three years that follow, providing 28,796 observations across the 7,035 firms.[18] I then limit the sample to include only firms that also meet the total assets and turnover eligibility requirements for higher grant rates in the year prior to winning a grant to take all three criteria into account while still relying on the binding criteria (employment) to define treatment status.[19] I omit the top 1% of the non-zero R&D expenditure distribution for firms with fewer than 100 employees (firms within an equal bandwidth around the threshold), as firm-level innovation investments can be highly volatile (Bronzini and Iachini 2014). The remaining data set contains 20,398 observations across 5,301 firms.

While the successful matching rate of 97% indicates that BvD's coverage is quite comprehensive, not all information is fully populated across all firms. Perhaps most importantly, of the 20,398 observations, only 4,703 observations (23%) across 1,614 firms (30.4%) have employment data for defining treatment status. My research designs implicitly condition on

[17] In Appendix A, I discuss the advantages of using FAME to study small firms relative the Business Enterprise Research and Development (BERD) data set that I use to study larger firms.

[18] This helps balance the number of post-grant years included for each award. Furthermore, some firms receive multiple grants and may be defined as treated when receiving their first grant but not defined as treated by the next grant (because of growing beyond the thresholds) or vice versa, and these funded years may overlap. Limiting it to a few years after receiving a grant helps limit the proportion of cases of when the treatment status changes while the firm may still be affected by the initial grant.

[19] If both variables are missing data, I assume the firm meets the criteria to preserve the sample size. I show in later sections that the results are robust to not conditioning on these criteria.

reporting employment, which helps mitigate some identification concerns, as otherwise one might be concerned that firms choosing to report variables like employment vary systematically from those that do not. At the same time, small firms in the UK can choose to report less detailed information in their P&L statements, which may introduce bias if firms under the threshold choosing to report other information (like R&D) differ systematically relative to the firms with populated data above the threshold. I carry out several tests in subsequent sections that suggest this does not seem to be an issue in my setting. The likelihood of reporting R&D and profits does not differ at the threshold and the main results hold when conditioning on reporting other financial variables.[20]

Table 2 provides descriptive statistics of grant awards and the main outcome of interest (R&D expenditures (£000s)). Column 1 provides information for firms of all sizes. In Column 2, I restrict the sample to firms within the MSE-optimal bandwidth around the 50-employee threshold (29 to 71 employees in the year before winning a grant), which I use as my baseline estimation sample.[21] It includes 1,356 observations across 417 firms that received a total of 561 grants between 2005 through 2017. The average award amount was £243k, covering about 53% of proposed project costs.[22] Firms reported investing an average of £154k. In Column 3, the bandwidth is widened to include firms with 1 to 100 employees, which have similar average award amounts and R&D expenditures but offers a much larger sample size.

4.2 Larger Firms

To study larger firms, I collect data from the UK's Business Enterprise Research and Development (BERD) database and Business Structure Database (BSD), which are confidential data sets created by the Office of National Statistics (ONS) accessed through the UK Data Service SecureLab. The BERD survey provides information on R&D expenditures for firms identified as actively performing R&D and is widely used to study larger firms in the UK.[23]

[20] See Section 5.4 and Section 6.3.

[21] I calculate the MSE-optimal bandwidths for the low and high tax credit rate periods following Calonico, Cattaneo and Titiunik (2014) and use the average throughout my baseline regressions. I show that the results are robust to much wider samples as well.

[22] Note that project costs and award amounts usually are spread over a few years whereas R&D expenditures are reported as an annual average.

[23] As I describe in detail in Appendix A, BERD data comes with many limitations for small firms but is suitable for firms that I call "larger." For example, see Guceri and Liu (2019).

Although firm size is reported in the BERD database, it is measured at the reporting unit level whereas tax credit rates are determined at the enterprise group level, so I gather employment data from the BSD, which covers this information for the universe of UK firms.[24] I aggregate the key variables to the enterprise group level to determine treatment and match firms over time to create an unbalanced panel from 2000 through 2014, with the pre-tax credit rate policy including years 2000 through 2008 and the post-tax credit rate policy period including years 2009 through 2014.[25] The final dataset consists of about 2,000 to 2,500 enterprise groups per year. A full discussion of the data sources, preparation, and matching procedures can be found in Appendix A.

The R&D expenditures measure in BERD is broken down by source of financing (i.e., the central government, internal finance, and external private finance). I proxy for "direct subsidies" with the amount that is funded by the central government. This can include grants allocated through funding competitions as well as other forms of direct support.[26] Appendix Table C.2 provides summary statistics of the final post-policy period (2009-2014) data set, which includes firms with 250 to 750 employees. One observation to highlight is, unsurprisingly, larger firms make more substantial R&D investments than small firms.

5 Validity of Research Designs for Small Firms

Interpreting the estimates from the approaches described in Section 3 as causal relies on several identifying assumptions: (i) the running variable is determined before treatment is assigned and is not precisely manipulated around the cutoffs, (ii) potential outcomes and determinants of outcomes are smooth across the threshold absent treatment, (iii) there are no other policies generating different incentives for firms to invests in R&D around the threshold, and (iv) the tax credit rate increase did not induce manipulative sorting or differences in potential outcomes around the grant generosity threshold. This section investigates these

[24]These data are derived from the Inter-Departmental Business Registrar (IDBR), which is a live register of administrative data collected by HMRC.The BSD includes all businesses that are liable for VAT and/or have at least one member of staff registered for the Pay as You Earn tax collection system.

[25]Like Dechezleprêtre et al. (forthcoming), start the post-policy period in 2009 because the new size thresholds in the R&D tax credit policy that I use in my empirical approach were not implemented until 2008, so 2009 was the first year in which firms with up to 500 employees were eligible for the higher rates.

[26]Importantly, the variable does not include funding received through R&D tax credits.

assumptions for small firms using the diagnostic tools that have become standard in the literature for RDDs, and for the difference-in-discontinuities research designs, I combine and extend tests commonly used to probe both RDDs and difference-in-differences.

5.1 No Running Variable Manipulation

I first explore whether there appears to be manipulative sorting (for the RDD) or *changes* in such sorting induced by the increases in tax credit rates (for the diff-in-disc). If particularly savvy firms strategically downsize to secure more generous funding, or if firms under the threshold intentionally limit their growth, firms just below the threshold may differ systematically, making firms just above the threshold a poor control group.

I follow the standard approach of examining the density of firms around the 50-employee grant generosity threshold for the entire sample period (2005-2017) and then separately for the "low tax credit rate" period (2005-2012) and "high tax credit rate" period (2013-2017). If firms manipulate size in response to the program design or policy changes, one would expect a bunching of firms just below the 50-employee threshold.

Figure 1 plots the distribution of firms by size using employment levels from the year before receiving a grant and fitting a third-order polynomial to the data. I use the baseline estimation sample (i.e., firms with 29 to 71 employees that also meet the turnover and total assets eligibility criteria). Panel A includes the entire sample period (2005 through 2017). There does not appear to be bunching under the threshold, and consistent with this, the formal McCrary density test does not detect a statistically significant discontinuity.[27] It also appears as though the *difference* in the densities between tax credit rate periods is continuous. That is, there is no evidence of a discontinuity in either time period (Panels B and C) and the difference in the discontinuity is also not statistically significant.[28]

[27] The discontinuity estimate (log difference in density height) at the 50 employee small firm threshold (and standard error) is 0.357 (0.335).

[28] The log differences (and standard errors) are 0.559 (0.623) in the low tax credit rate period and 0.300 (0.372) in the high tax credit rate period. Applying a simple t-test for statistical significance in their difference provides a t-statistic of 0.357

5.2 Balance in Pre-Determined Characteristics and Covariates

The RDD estimates are only unbiased if potential outcomes are balanced, and similarly, diff-in-disc research designs require there to be continuity in the differences in potential outcomes. To explore whether this is likely, I test whether there are discontinuities and differences in discontinuities in pre-treatment R&D and other observable covariates that could determine firms' R&D investments. This also provides an indirect test of manipulative sorting to supplement the bunching analysis above; if firms manipulate their size and those that do are systematically more savvy, they may also perform better on other financial indicators.

To carry this out, I restrict the sample to observations prior to when firms receive their first grant and estimate the models of Equation 1 and 2 using R&D, (lagged) cumulative R&D, age, (log) total assets, and (log) current liabilities as dependent variables.[29] Appendix Table C.3 reports the results. Firms appear similar around the threshold according to these measures and the increase in the tax credit rate generally does not induce discontinuities. One exception is that there is a statistically significant discontinuity and difference-in-discontinuities in current liabilities (at the 10% level). The magnitudes of some of the other coefficients are also non-trivial even though they are not statistically different than zero. This motivates me to include these variables as controls in my baseline specifications throughout the paper to address pre-grant differences between treated and untreated firms.

5.3 Selection and Preferential Treatment of Small Firms

Two related potential concerns about the research designs I employ include selection of firms applying for grants and preferential treatment of small firms by the funding agency. If firms applying for and winning grants are already more innovative than those that do not apply or win, then a comparison of the effect of grant funding on grant recipients relative to non-recipients could be biased by unobserved characteristics determining innovative capacity. While my estimation approaches rely primarily on a comparison of outcomes only for grant recipients (as opposed to non-recipients) by examining the effect of different grant rates on

[29] To be consistent with the approach taken throughout the main analyses, the sample includes firms in the MSE-optimal bandwidth (29 to 71 employees using lagged employment) that also meet the total assets or turnover eligibility requirements. All running variable controls are included in the regressions.

winners' outcomes, which alleviates some of these concerns, selection may still be problematic if the selection itself differs systematically at the threshold. For example, if firms under the size threshold are more likely to apply because the grant funding generosity rates are higher, the quality of projects proposed or firms' sensitivities to R&D cost shocks may differ.

Second, if the funding agency preferentially treats certain firm size groups, there may be underlying differences in the quality of projects funded for firms just under the thresholds relative to over. While Innovate UK most frequently includes small- and medium-sized enterprises (SMEs) rather than just small firms when setting eligibility requirements for competitions that target smaller organizations, there could exist broader unobserved government or agency objectives to support small firms in particular.

Although I do not have data on all applicants to test whether there are differences in application quality or applicants' observable characteristics, there are a few indirect pieces of evidence that suggest these probably are not significant threats to my identification strategies. Consider the results from the covariate balance tests in the preceding sub-section. Although this is not a comparison of grant winners and losers, seeing that there are few differences in observables around the threshold for firms that do win grants at some point but in their *pre*-grant years suggests that firms around the threshold were similar pre-treatment. Of course, it could still be that all firms receiving grants overall are stronger before winning a grant relative to those that never win, but in my context of focusing on firms that do win grants, what is most important is that there are no major differences for firms just below and above the threshold. If firms were systematically different in their innovative capacity below and above the threshold or if Innovate UK preferentially treated small firms, the balance tests likely would have detected discontinuities in variables like cumulative R&D.

I also directly examine whether the likelihood of winning a grant is smooth across the threshold for these firms. The sample is still limited to firms that do eventually win, but if discontinuities in the likelihood of winning a grant for this set of firms are detected, this could suggest that firms under the threshold are indeed systematically different in their innovative capacities or are favored in the evaluation process. For this estimation, I include all years whether the firm has received a grant or not and use an indicator equal to one if the firm receives a grant that year as the dependent variable. Appendix Table C.4 provides the RDD results in Columns 1-3 and diff-in-disc results in Columns 4-6 when including different sets of controls and limiting the sample in various ways. The estimates suggest there are no discontinuities or differences in discontinuities in the likelihood of receiving a grant.

5.4 Other Policies or Policy Changes

The final assumptions are that there are no other policies generating different incentives to invest in R&D at the 50-employee threshold and no other policy changes at the same time as the shift from a low to high tax credit rate period that would induce such differences. In a sense, observing no firm size manipulation or discontinuities in pre-treatment covariates is consistent with this being true, as confounding policies would likely induce similar sorting. However, I manually reviewed many UK programs and policies with size-based preferential treatment (see Appendix Table C.5 for descriptions).[30] The majority target both small- and medium-sized enterprises as opposed to only small firms, creating different incentives at the 250-employee threshold rather than 50. Those that do preferentially treat smaller firms tend to define "small" using other criteria.

One potential threat, though, is that small firms faced different reporting requirements through my study period and could choose to file less detailed accounts. While BvD conducts significant additional research to fill these gaps, coverage of some key variables could still differ around the 50-employee threshold either because of reporting itself or BvD's coverage. This could bias the results in either direction if firms that choose to report versus those that do not are systematically different. If firms that report are systematically stronger financially, for example, they may be more likely to apply for and win Innovate UK grants, but they also may be less sensitive to cost shocks.

The balance tests likely would detect discontinuities in pre-treatment observables if this is a major issue. Furthermore, my baseline sample inherently conditions on reporting employment given that it is the running variable, so it is more likely that missing R&D data represent true zeros given that these firms choose to report this information. Nonetheless, I

[30] Policies in other countries have been shown to induce bunching around small firm thresholds, for example. In France, many labor laws start to bind for firms with more than 50 employees, and this results in bunching just below the 50 employee threshold (Garicano, Lelarge and Van Reenen 2016).

investigate this further by directly examining whether there is a discontinuity in the likelihood of reporting profits and R&D at the threshold in the years prior to when firms receive their first grant.[31] Appendix C.6 provides the results from estimating the RDD model using indicator variables equal to one if the variable is not missing data and zero otherwise. There are no statistically significant discontinuities, suggesting that there are no differences in either reporting or coverage of these variables in FAME for this set of firms.

Finally, the UK's introduction of a patent box policy in 2013—which provides corporate tax relief on profits earned from patents—could confound the subsidy interaction estimates if it also enhances the effect of grant funding on R&D. I address this in Section 6.3 and conclude that it is an unlikely confounder based on the policy's design, statistics on its uptake so far, existing studies of patent box policies, and additional robustness checks.

6 Main Results

6.1 Effect of Grant Funding on Small Firms' R&D

To examine the independent effect of more generous grant funding on its own, I start by plotting mean R&D expenditures for bins of firms using the baseline sample in Figure 2. Panel A includes observations for years prior to when firms receive their first grant, and as expected, R&D appears to be relatively smooth across the threshold. On the other hand, for the years in which firms receive grants and the three years that follow, a discontinuity at the treatment threshold emerges.

The econometric results align with these plots. Panel A of Table 3 presents the findings from estimating the RDD model of Equation 1 for the baseline sample. Column 1 provides estimates from when not including any controls or fixed effects, which suggest that R&D expenditures are about £195k higher for firms just under the 50-employee threshold relative to those just over it. The coefficient increases slightly and remains statistically significant when gradually adding different sets of fixed effects and controls.[32] In the final specification of

[31] I only include pre-grant years because missing data may simply reflect zero expenditures and receiving a grant may make it more likely that R&D is non-missing in the data because they did actually invest, thus capturing the policy effect of interest as opposed to differences in reporting.

[32] Following Dechezleprêtre et al. (forthcoming) and Howell, Rathje, Van Reenen and Wong (2023), I

Column 5 when including all controls and fixed effects, I find about a £202k difference in R&D at the threshold. This is a little more than double the sample's mean R&D expenditures.

I also probe the continuity in potential outcomes assumption of the RDD further by estimating this same set of regressions using observations for years prior to when firms received their first grant and do not detect any statistically significant discontinuities (Panel B). The magnitudes of the coefficients are also small. To provide further confidence that the discontinuities are not just an artifact of the data, I conduct placebo tests by estimating whether there are effects at "fake" cutoff points and do not detect any positive discontinuities (see Appendix Table C.7).[33] This holds when using triangular weights and an even bandwidth (Panel A) as well as uniform weights and the baseline bandwidth of 21 employees when possible (Panel B).[34] The coefficient estimates are small in magnitude and are always negative, if anything, and the only case for which the estimate is statistically significant is when using a 10 employee threshold in Panel B (but with a negative sign).

6.2 Effects of Policy Interactions on Small Firms

6.2.1 Graphical Exposition of Discontinuities

Turning to the policy interactions, I start by plotting average R&D for bins of firms in Figure 3 in four separate plots that are consistent with the econometric estimates that follow. Panel A includes observations for years prior to when firms receive their first grant in the low tax credit rate period (2005-2012) and Panel B includes observations from the baseline estimation sample (when firms receive grants and the three years that follow) in the low tax credit rate period. As expected, there is no indication of a discontinuity in Panel A, but once firms receive a grant, the slope of the relationship for firms under the grant rate threshold appears

control for cumulative pre-award R&D expenditures and other covariates to help ensure that the results are not driven by pre-existing differences between the treatment and control groups. The full set of "additional controls" includes the interaction and main effects of lagged cumulative R&D expenditures and firm age, (log) total assets, and (log) current liabilities. In Panel A, the interaction and main effects of R&D prior to winning a grant interacted and the 50-employee threshold indicator are also included.

[33] I use thresholds that capture relatively small firms but which are far enough from the 50-employee threshold to not be contaminated by the actual treatment effects (5, 10, 15, 85, and 90 employees).

[34] Specifically, in Panel A, I use the baseline bandwidth of 21 on each side when possible, but for lower thresholds (e.g., 10, 15, etc.), I use the largest bandwidth possible while keeping it even. For example, for the 15-employee threshold, the sample includes firms with 1 to 29 employees. In Panel B, the baseline bandwidth of 21 is used whenever possible even if that leads to unequal bandwidths on each side of the threshold.

to pivot upwards (Panel B). The change is subtle but apparent.

Panels C and D of Figure 3 plot R&D for bins of firms using data from the high tax credit rate period (2013-2017), with observations prior to when firms receive their first grant in Panel C and observations from the baseline estimation sample in Panel D. Once again, there is no clear discontinuity before firms receive grants, but once they do, there is a significant increase in the slope of the line for firms under the 50-employee threshold such that a large discontinuity emerges at the threshold. This begins to suggest that higher grant funding rates increased R&D by more when tax credit rates were also higher.

6.2.2 Main Econometric Results

Moving to the econometric analysis, I estimate the policy interaction effects on R&D using the model of Equation 2 and present the results in Table 4. The diff-in-disc estimates are in the first row, reflecting the change in the discontinuity in R&D between the low and high tax credit rate periods. The findings indicate that, for firms receiving Innovate UK grants, the marginal effect of more generous grant funding rates increases substantially when tax credit rates increase, suggesting that the two instruments are complements.

In Column 1, no controls or fixed effects are included and I find that the discontinuity increases by £311,000. I then gradually add controls, which are not required for identification but can improve precision and help address pre-existing differences between firms under and over the threshold should they exist.[35] Column 2 includes firm and year fixed effects and the estimate increases to £436k. In Column 3, I add variables related to firms' previous R&D efforts—the firm's average R&D in years prior to when it receives its first grant interacted with the treatment dummy so that the effect can differ on each side of the threshold as well as firm age interacted with lagged cumulative R&D (along with all main effects)—and I include (logged) total assets and current liabilities in Column 4. The magnitudes of the coefficients increase to £416k and £445k, respectively, and the estimates are statistically significant at the 5% level. Lastly, I include industry-year fixed effects in Column 5 to account for how technological opportunities and government funding preferences may change differently over

[35] While there was only one statistically significant difference detected when carrying out the balance tests of pre-determined covariates in Appendix Table C.3, the magnitudes of the differences in some cases were non-trivial.

time across industries and the magnitude of the estimated effect increases to £560k.

I estimate this same set of regressions for years prior to when firms receive their first grant to probe whether these discontinuities emerged by chance. If macroeconomic trends or changing government preferences differentially impact firms just below and above the threshold, differences in discontinuities may arise due to other factors. I should detect such effects in pre-grant years as well if this is the case but there are indeed no differences in the discontinuity when firms have not yet received a grant (see Appendix Table C.8).

Finally, I estimate an event study version of the diff-in-disc model by interacting the treatment variables with an indicator for each year (using 2012 as the reference period).[36] The annual coefficients are plotted in Appendix Figure B.1 with their 95% confidence intervals included in Panel A, and for ease of interpreting the magnitudes of the coefficients, I omit the confidence intervals in Panel B. The point estimates for the discontinuity are positive (as expected) in the low tax credit rate period (but not statistically significant due to having small sample sizes for each year). However, the discontinuity increases sharply in 2013 when shifting to the high tax credit rate period, consistent with the main econometric results. The coefficient estimate averages are 270.7 and 536.6 in the low and high tax credit rate periods, respectively (see Appendix Table C.9).[37] The event study also provides an indirect test of the parallel trends assumption. In my setting, the assumption is that the discontinuity in R&D at the threshold would have followed the same trend over time in the absence of the tax credit rate increase.[38] Indeed, while R&D expenditures are higher for firms just under the grant generosity threshold in earlier years (as expected), the discontinuity is stable.

6.2.3 Interpretation of Magnitudes

The magnitudes of the difference-in-discontinuity estimates are economically significant. The baseline estimate of £560k is triple the average level of R&D expenditures for this sample of firms over the full time period and 2.8 times the discontinuity found when estimating the independent effect of more generous grant funding on its own.

[36] I also interact the running variable controls and include all baseline fixed effects and controls.

[37] Appendix Table C.9 provides the coefficients and their standard errors. Note that 2015 appears to be an outlier year in which the discontinuity drops significantly, but the magnitudes of the point estimates in other high tax credit rate years are consistently higher than those in the low tax credit rate period.

[38] This is different than testing for discontinuities in R&D in pre-grant years, which I do in Table 3.

One potential explanation is that the combination of policies alleviates financial constraints. This would be consistent with other papers that find particularly high effects of grants and tax credits for small firms even on their own, with alleviating liquidity constraints being a common explanation (Howell 2017; Dechezleprêtre et al. forthcoming).[39] If the combination of subsidies enables firms to start new projects that they previously could not pursue due to high upfront costs, or if non-R&D performers start to invest for the first time, the change in R&D might be especially high because of high start-up costs. As pointed out by Mohnen and Röller (2005), a package of policies (rather than just one) may be required to increase the propensity to invest in R&D whereas a single policy may induce more R&D on the intensive margin. I provide evidence consistent with large positive interaction effects alleviating financial constraints in Section 7, including through an extensive margin effect.

6.2.4 Widening the Firm Size Window

To explore whether the effects are restricted to firms only within a narrow window around the threshold, I estimate the effects using wider ranges of firm size (see Appendix Table C.10). I first extend the bandwidth under the threshold to see whether much smaller firms are sensitive to the grant rate difference (Columns 1-3) and the results are similar to the baseline.[40] When expanding the window to include much larger firms (Columns 4-7), the magnitudes of the coefficients are also remarkably stable. The difference in discontinuity is £529k when including firms with 10 to 200 employees and decreases to £471k once widening to 250 employees.[41] These results suggest that the positive interaction effects are not driven by sample choice and may apply to a wider range of firms.

[39] Others finding high effects of grant funding for small firms include Lach (2002) and Howell et al. (2023).

[40] The diff-in-disc estimate is £531k and statistically significant at the 1% level when including firms with 10 to 71 employees, for example.

[41] I stop at 250 because this is the threshold at which many other policies in the UK bind, as it is the employment cutoff determining SME status for most other policies and programs.

6.3 Falsification Tests and Robustness Checks

6.3.1 No Evidence of Relabelling

Reported R&D might increase without reflecting changes in innovative activity if firms simply relabel ordinary investments as R&D to reap larger rewards (Hall and Van Reenen 2000; Agrawal et al. 2020).[42] While relabeling is less of a concern for grants given how project finances associated with Innovate UK projects are closely monitored, government oversight of spending under tax credit schemes can be more difficult and costly. I estimate the effects on three measures of non-R&D investment (see Appendix Table C.11), which should decrease if firms relabel ordinary investment as R&D: tangible assets (in levels) (Column 1), the change in tangible assets plus depreciation following Zetlin-Jones and Shourideh (2017)'s measure of capital expenditures (Column 2), and the change in tangible assets while ignoring depreciation to avoid losing observations due to missing data (Column 3). There are no statistically significant effects and the magnitudes of the coefficients are small.[43]

6.3.2 Placebo Tests for Grant Rate Threshold

Like before when examining the effect of grant funding on its own, I perform placebo tests by estimating whether there are differences in discontinuities using pseudo-thresholds across which expenditures should be smooth. The results are presented in Appendix Table C.12 using triangular weights and even bandwidths in Panel A and uniform weights and the baseline bandwidth in Panel B.[44] None of the estimates are statistically significant and most coefficients have negative signs, if anything. For the three exceptions when the sign is positive, the magnitudes of the coefficients are small relative to the baseline estimates.

[42]This has been shown to occur in other settings. For example, Chen, Liu, Suarez Serrato and Yi Xu (2021) find that firms in China engaged in relabeling in response to corporate tax rule changes.

[43]Since the tangible assets and investment variables are highly skewed, I follow a similar trimming rule as I do for R&D in the baseline by omitting the top 1% of the baseline estimating sample's distribution. The results remain statistically insignificant when not omitting any outliers as well, though.

[44]Specifically, in Panel A, I use the baseline bandwidth of 21 on each side when possible, but for lower thresholds (e.g., 10, 15, etc.), I use the largest bandwidth possible while keeping it even. For example, for the 15-employee threshold, the sample includes firms with 1 to 29 employees. In Panel B, the baseline bandwidth of 21 is used whenever possible even if that leads to unequal bandwidths on each side of the threshold.

6.3.3 Placebo Test for Policy Change Timing

I also carry out placebo tests for the tax credit rate change timing by testing whether there are diff-in-disc effects when assuming tax credit rates go from low to high in other years. I estimate separate equations that impose each year as "pseudo" policy change years and plot the results in Appendix Figure B.2.[45] The only case in which there is a large, positive, and statistically significant diff-in-disc effect is for the actual treatment year (2013), as expected if the effects are not driven by other changing conditions throughout the sample period. All other coefficients are much closer to zero and statistically insignificant.

6.3.4 Other Policy Changes in 2013

One potential concern is that the UK's Patent Box policy, introduced in 2013, could confound the results if it also enhances the impact of grant funding. Experts have argued that it is poorly targeted for promoting research (Griffith, Miller and O'Connell 2010), though, consistent with evidence in the literature on patent box policies as well. While patent box regimes reduce patent transfers out of countries, they do not seem to significantly affect R&D (or patenting levels) (Gaessler, Hall and Harhoff 2021; Alstadsaeter, Barrios, Nicodeme, Skonieczna and Vezzani 2018).[46] See Hall (2022) for a comprehensive review.

Second, data on patent box and R&D tax credit claims suggest that the latter are much more important for small firms. I gathered data on the number and value of claims for both policies from the HMRC's official public statistics (see Appendix Table C.13). On average, between fiscal years 2013/14 and 2016/17, small firms filed only 250 patent box claims per year on average (totaling £6.3 million per year). To make the statistics comparable to tax credit claims, which are provided for SMEs as a whole, I also gather data on medium-sized firms, which filed 249 patent box claims per year (£21.7 million per year) for a sum of 499 patent box claims by SMEs per year. In contrast, SMEs made an average of 31,976 R&D tax relief claims per year (totaling £1,511 million per year) over this time period, which is about

[45] For example, the 2007 coefficient is associated with assuming that tax credit rates enter a high rate period in 2007 rather than 2013.

[46] One exception is Mohnen, Vankan and Verspagen (2017), who find that the patent box policy in the Netherlands is associated with more R&D person-hours. However, the Netherlands' policy uniquely covers non-patentable R&D as well.

64 times the number of annual patent box claims. One complication with this comparison is that SMEs under the tax credit policy include firms up to 500 employees whereas the cutoff is 250 for the patent box policy, but even if small firms account for only 10% of all R&D tax credit claims, this still amounts to 12.8 times more than patent box claims by small firms.

Results from two empirical tests also provide assurance that the patent box policy is probably not driving the results. First, omitting firms in the manufacturing and wholesale and trade industries—which make up more than 50% of patent claims each year—does not weaken the findings (see Appendix Table C.14).[47] If the patent box policy was a major issue in my setting, one might expect a lower or non-existent policy interaction effect for firms in industries that do not make many claims. Second, one might expect firms with a longer track record of investing in R&D to benefit most from the patent box policy, as they are more likely to already own more patents or be closer to filing patents. The baseline regressions already control for cumulative R&D, but as shown later in Section 7.1.2, the effects are particularly substantial for firms with *zero* cumulative R&D if anything (Column 2 of Table 6).

6.3.5 Distribution of R&D and Missing Data Assumption

Firm-level R&D tends to be highly skewed and can also have many zeros (see Appendix Figure B.3 for the distribution of R&D in my sample).[48] To ensure the results are not sensitive to the winsorization rule I apply in the baseline, I skim the top 5% of non-zero R&D (rather than 1%) and the results are similar (see Column 1 of Appendix Table C.15).[49] I also use normalized versions of R&D as the dependent variable—R&D per employee and R&D as a proportion of total R&D in its 4-digit SIC—and the effects become stronger.[50] Lastly, since I assume that missing R&D data are zeros but some financial data may be less frequently reported for small firms (so the likelihood that missing R&D data are truly zeros may differ at the threshold), I estimate the baseline model conditional on firms also reporting profits (Column 4) and cost of sales (Column 5). The effects again become stronger.

[47] Firms with SICs of 10 to 33 (manufacturing) and 45 to 47 (wholesale and trade) are excluded.

[48] The 75th percentile of the distribution is zero and the 99th percentile (after winsorizing) is £3.2 million.

[49] The diff-in-disc estimate decreases to £508k and remains statistically significant.

[50] The diff-in-disc estimate for R&D per employee is 10.6, more than triple the sample's mean value of 3.38, and the effect on R&D as a proportion of industry total is 0.11 relative to a mean of 0.018.

6.3.6 Additional Robustness Checks

The results are robust to various other tests of my sample selection and modeling choices. In Column 1 of Appendix Table C.16, I use the MSE-optimal bandwidth for the entire sample time period overall rather than averaging the MSE-optimal bandwidths for high and low tax credit rate periods.[51] In Column 2, I return to using the baseline bandwidth but do not condition on firms also meeting the other grant generosity rate criteria in case doing so introduced bias associated with missing data. In Column 3, I omit observations associated with grants received after 2015 so that there are at least 2 years of post-grant data for all grants, and in Column 4, I include four years of post-grant data rather than three. The estimates remain higher than £500k and statistically significant in these four regressions. In Columns 5-7, since firms may have been particularly constrained and cost-sensitive during the Great Depression, I omit years prior to 2008, 2009, and 2010 and the effects become stronger, if anything. Lastly, the results do not appear to be sensitive to my clustering, weighting, and running variable flexibility decisions (see Appendix Table C.17).

7 What Drives Subsidy Complementarity?

The remainder of this paper focuses on exploring what drives the large, positive subsidy interaction effects. There are four main potential explanations: (i) alleviating financial constraints, (ii) changing government preferences, (iii) substitution away from applying for grants, and (iv) cumulative grant funding. I provide evidence that makes it hard to reconcile (ii), (iii), or (iv) as the underlying mechanisms. Rather, the evidence is most consistent with the combination of policies helping firms overcome financial constraints.

7.1 Financial Constraints as the Key Mechanism

In this section, I provide three sets of evidence that are consistent with financial constraints as the underlying driver of subsidy complementarity: 1) the interaction effects are much larger for firms that appear to have been more constrained in pre-grant years, 2) there are

[51] The bandwidth in this case is 19 rather than 21.

large extensive margin effects whereby the subsidy interaction induces "R&D entry" and increases the propensity to invest in R&D for previous R&D performers as well, and 3) the subsidies appear to be substitutes for larger firms, which tend to be less constrained.

7.1.1 Heterogeneity in Small Firm Effects by Financial Constraint Proxies

I first test whether there is heterogeneity in the policy interaction effects based on whether firms appear to have been constrained before receiving grants, as the effects should be larger for constrained firms if the policies enabled firms to overcome the high upfront costs of starting a new project. I use three proxies for constraints that aim to capture the amount of funds firms may have available to self-finance R&D: short-term debt, operating profit, and an "available funds" measure that I construct. Firms facing financing constraints are more likely to have short-term loans and overdrafts (i.e., be above the median (absolute) value of short-term debt), as these types of loans are typically sought when firms do not have internal funds to cover unexpected costs. Firms are also more likely to be able to internally-finance projects when their operating profits are healthier. Lastly, I calculate "available funds" following the approach of Zetlin-Jones and Shourideh (2017) as the sum of before-tax profits and depreciation, capturing internal resources available for investment.[52]

I split the sample using the median values of these proxies in the year before firms in the baseline sample receive grants and estimate the policy interactions separately for firms that appear constrained versus unconstrained. The results are in Table 5 with the effects on constrained firms in odd-numbered columns and effects on unconstrained firms in even-numbered columns. The estimates are positive and statistically significant only for constrained firms and the magnitudes of the coefficients are large. Firms overcoming constraints also helps explain the large interaction effects in the baseline results, as the effects on constrained firms are pulling up the average effect.

7.1.2 Extensive Margin Effects

It is possible that firms with no previous R&D wanted to pursue projects but did not have the resources to finance building a new lab or purchasing machinery. Similarly, firms that did

[52] See Appendix Table C.18 for more detail on these measures.

previously perform R&D also may not invest in some years despite having viable projects if they are constrained. As such, if the combination of policies induces R&D "entry" or increases the *propensity* of firms to invest, this could signal that the availability of both of subsidies alleviates financial constraints. Extensive margin effects like these also may lead to especially substantial effects on R&D as firms go from zero investment to some large positive investment associated with starting a new project.

To explore this, I first estimate the diff-in-disc model using an indicator variable equal to one when R&D expenditures are positive and zero otherwise as the dependent variable, limiting the sample to only observations for which there is no cumulative R&D (in year $t-1$). The results in Column 1 of Table 6 (Panel A) indicate that the policy interaction indeed increases R&D entry by 50%. Furthermore, in Column 2 of Panel A, I estimate the policy effects on R&D conditional on having no previous R&D and find that the interaction doubles relative to the baseline estimates. These results suggest that R&D entry likely plays a role in explaining the large magnitude of the policy interaction effects.

I also examine whether the policy interactions increase the propensity to invest without conditioning on having no previous R&D investments. The coefficient estimate is not statistically significant when using the baseline bandwidth (Column 3 of Panel A), but once widening the window around the threshold to increase statistical power, the estimate increases and becomes statistically significant (Columns 4-5). The combination of policies appears to increase the propensity to invest in R&D by about 27-29%.

Lastly, I carry out the same analyses for the independent effect of grant funding to investigate whether these effects are primarily driven by grant funding as opposed to policy interactions. The results are presented in Panel B of Table 6. I find no effect of more generous grant rates on R&D entry or the propensity to invest on their own. At the same time, the effect on R&D expenditures (£198k) for these firms is nearly identical to the baseline estimates. These findings suggest that the grant generosity effects on R&D might be largely driven by firms that were particularly constrained (Column 2, Panel B) but it is the combination of policies that helps them "enter" (Column 1, Panel A). These results are consistent with Mohnen and Röller (2005)'s finding that a package of innovation policies can increase the propensity to innovate.[53]

7.1.3 No Evidence of Complementarity for Larger Firms

Lastly, I estimate the policy effects for larger firms, which tend to be less financially constrained than small firms, so finding no complementarity would be consistent with constraints playing a role. I first estimate the independent effect of being eligible for more generous tax credits and then the interaction with direct subsidy funding following the research designs detailed in Section 3. All four approaches primarily leverage the discontinuity in tax credit rates, whereby firms under a 500-employee threshold benefit from being eligible for much more generous tax credits than those over the threshold.[54]

Identification Assumptions.— The identification assumptions are analogous to those for the small firm analyses but apply to the 500-employee tax credit rate generosity threshold rather than the 50-employee grant rate threshold. The first is that there are no other policies generating different incentives for firms around the threshold. The R&D tax relief scheme is the only policy in the UK for which a 500-employee threshold is used to define SMEs, with most policies using the standard 250 employee cutoff for defining SME eligibility.

Second, firms must not select into higher tax credit rates, such as by strategically downsizing. I plot the firm size distribution densities for the pre-policy period (2000-2008) in Panel A of Appendix Figure B.4 and the post-policy period (2009-2014) in Panel B for firms with 250 to 750 employees, the baseline sample window that I use. There do not appear to be discontinuities across the threshold in either period visually and the formal McCrary tests do not detect bunching. The difference between the differences at the thresholds in the pre- and post-policy periods is also not statistically significant.[55]

Third, the cutoff determining whether firms quality for higher tax credit rates must not be endogeneously determined by firm characteristics. I test whether there is continuity in firm

[53] Note that this is not to say that receiving a grant as a whole does not have an effect like this on the extensive margin. Rather, the higher *rates* on their own do not appear to do so.

[54] Note that there are also total asset and turnover thresholds applied by the policy for determining SME status under the R&D tax relief scheme, but since I use employment for the estimation, I will mostly refer only to the 500-employee threshold.

[55] The McCrary density tests provide discontinuity estimates (log differences in density height) at the firm threshold (and standard errors) of -0.063 (0.284) in Panel A and -0.141 (0.289) in Panel B.

characteristics and covariates—particularly those that may be related to R&D effort—around the threshold in pre-policy years (2000-2008) by estimating a RDD model and using the following variables as outcomes: direct subsidies for R&D (£000s), the proportion of R&D expenditures supported by direct subsidies, revenue (millions), labor productivity (revenue per employee), age, average R&D worker wages, and the number of R&D scientists (see Appendix Table C.19).[56] There are no statistically significant discontinuities, suggesting that firms just below the threshold are similar to those just above it on these measures before a difference in tax credit rates existed at this cutoff. As with the small firm diagnostics, these balance tests also indirectly suggest there is no evidence of manipulative sorting.

Lastly, although the causal effect of direct subsidy funding cannot be identified for larger firms, I interpret the *discontinuity* in the correlation between direct subsidies and R&D as the policy interaction effect since it is driven strictly by the exogenous tax credit rate threshold. The assumption is that the endogeneity of direct subsidies goes in the same direction and is a similar magnitude for firms just below and above the threshold. While this cannot be tested directly, the covariate balance tests in Appendix Table C.19 are consistent with this holding, as firms would likely differ on observables otherwise. Consider how a primary endogeneity concern is that more innovative and R&D-intensive firms could be more likely to apply for and win grants. Firms just under and over the threshold invested similarly and received similar amounts of direct funding prior to the tax credit policy implementation, though.

Independent Effect of Tax Credits for Larger Firms.—I now estimate the independent effect of being eligible for more generous tax credits following the approaches described in Section 3. When implementing the RDD and using data from post-policy years (2009-2014), I find that R&D expenditures are about £1 million higher for firms just below the cutoff relative to those just over it (Column 1 of Table 7).[57] This is a 57% increase in R&D relative to the pre-policy mean.[58] The result is similar but a bit lower than Dechezleprêtre et al. (forthcoming)'s baseline estimate of 86% when they use an RDD and administrative data.[59]

[56] As in the baseline regressions, I use triangular weights and include the running variable controls. I winsorize the sample by omitting the top 1% of the pre-policy R&D expenditure distribution to be consistent with the approach I take when estimating the policy effects. Standard errors are clustered at the firm level.

[57] Firms' average pre-policy R&D is also included as a control following Dechezleprêtre et al. (forthcoming).

[58] The pre-policy mean R&D for the baseline sample is £1.75m (see Column 1 of Appendix Table C.2).

[59] One potential driver of the slight difference could be that I use employment as the running variable to

I also examine whether there are discontinuities in pre-policy years (2000-2008) as a falsification test. As expected, there is no statistically significant correlation (Column 2). However, similar to Dechezleprêtre et al. (forthcoming), the magnitude of the coefficient is non-trivial (£191k). I therefore also estimate a diff-in-disc model to capture how the discontinuity changes between the pre- and post-policy periods and find that R&D increases by £539k-563k (Columns 3 and 4 of Table 7) for firms under the threshold. These estimates are nearly identical to those in Guceri and Liu (2019), who use the same data (BERD) and employ a similar method (a difference-in-differences approach using the employment threshold and pre/post variation). My estimates imply that R&D expenditures increase by 32% (Column 4) relative to the pre-policy mean while Guceri and Liu (2019) find an increase of 33%. Lastly, when widening the sample further to include firms with up to 1,000 employees and estimating the diff-in-disc model using uniform weights rather than triangular—mimicking more of a difference-in-differences approach—I find that R&D increases by 33.6%.

These results are robust to both narrowing and widening the bandwidth around the threshold, with statistically significant discontinuities in R&D ranging from 43% to 58% higher than the pre-policy mean when estimating the effects using the RDD (see Panel A of Appendix Table C.20). There are also no statistically significant discontinuities in the pre-policy period for these samples (Panel B).

Policy Interaction Effects for Larger Firms.—Table 8 provides the subsidy interaction effects for larger firms. In Columns 1-4, I estimate the discontinuity-in-effects model using data for the post-tax credit policy period (2009-14) and find that, while there is a positive and statistically significant correlation between direct subsidies and R&D, it is substantially dampened for firms just below the tax credit generosity threshold. In Column 2, when controlling for average pre-policy R&D, the correlation between direct subsidies and R&D decreases from 3.85 for firms just above the threshold to 1.02 for firms just below the threshold. Once including all controls and fixed effects (Column 4), the positive correlation is entirely diminished. On the other hand, I detect no dampening effects in pre-policy years, as expected (see Columns 5 and 6 of Table 8). There is still a positive and statistically signif-

determine treatment (while conditioning on the other two criteria), as employment is the binding criteria, whereas they use total assets.

icant correlation between direct subsidies and R&D but no discontinuity in this correlation, suggesting that the dampening effect found at the threshold in Columns 1-4 did not arise by chance or due to pre-existing differences.

Finally, I estimate the difference-in-discontinuities-in-effects model to test how the discontinuity in the correlation between R&D and direct subsidies changes between the pre- and post-policy periods using data across the full sample period (Column 7). As expected, while there is a positive correlation between direct subsidies and R&D in the pre-policy period but without a discontinuity (the coefficients in the top two rows of Column 7), the correlation is dampened significantly by the tax credit policy in the post-policy period.[60]

The results survive a number of falsification tests and robustness checks. First, I test whether the subsidy funding effect is smooth across arbitrary pseudo-thresholds where there is no difference in the tax credit rates and find no interaction effects, as expected (Columns 1-4 of Appendix Table C.21). As a second falsification test, I estimate the effects separately for non-capital and capital R&D expenditures. Most capital investments do not qualify for tax credits in the UK, so the substitution effect should primarily occur through non-capital expenditures (e.g., R&D labor) and this is indeed what I find (see Columns 5 and 6). The interaction effects on non-capital R&D are very similar to the baseline, but for capital expenditures, there is no statistically significant dampening of the otherwise positive relationship between direct subsidy funding and capital R&D expenditures.

Since I use only the current year's employment to define tax credit treatment status in the baseline but eligibility formally requires firms to fall under the thresholds for two consecutive years, I estimate the effects when defining treatment based on the current year plus the preceding year as well as the current year plus the preceding two years and the results are very similar (see Appendix Table C.22). Lastly, the estimates in Panel A of Appendix Table C.23 show that the discontinuity-in-effects estimates remain stable when increasing the flexibility of the running variable controls (Columns 1 and 2), using uniform kernel weighting rather than triangular (Column 3), and limiting the sample to include only observations for which there is a positive value of direct subsidies. In Panel B, I run the same

[60] The coefficient associated with the interaction is -1.94 and statistically significant at the 5% level. There is also still a large discontinuity in R&D detected in response to the tax credit policy on its own.

regressions for pre-policy years, and while the correlation between R&D and direct subsidies remains positive and statistically significant, there are no discontinuities in the correlation.

7.2 Unlikely Drivers of Positive Interaction Effects for Small Firms

7.2.1 Shift in Types of Projects Funded

Funding agencies often are given a fixed budget from the government and then have control over the focus and design of funding competitions, which is the case for Innovate UK. This opens an opportunity to favor certain industries, technologies, or research. While I control for how general trends in government preferences and technological opportunities might impact outcomes with industry-year fixed effects, preferential treatment also could occur within-competition, which could put upward pressure on the policy interaction estimates if it results in a shift in the mix of R&D that qualifies for tax credits (e.g., more labor-intensive projects). For example, while Innovation Leads typically recommend the highest-ranked project for funding, they may also take a portfolio approach to meet the competition's objectives (if stated in the competition description) and recommend projects that could complement others even if they are lower-ranked (InnovateUK n.d.). Firm behavior also might shift the types of R&D projects that are funded, as firms could propose projects that involve more R&D qualifying for tax credits when tax credits are more generous.

To investigate this, I estimate the policy interaction effects on indicators for whether projects fall within different categories of R&D activity (see Appendix Table C.24). There are no statistically significant effects on the likelihood that a funded project is a feasibility study (Column 1), proof of concept (Column 2), development of a prototype (Column 3), or proof of market (Column 4). Furthermore, when estimating the baseline model and controlling for within-competition unobserved factors with competition-level fixed effects (Column 5), the main estimates increase, suggesting that any within-competition favoritism that does exist is putting downward pressure on the results, if anything.

7.2.2 Substitution Away from Applying for Grants

Relatedly, firms may substitute away from applying for grants when tax credit generosity increases, as the benefits of tax credits may exceed the benefits of grants minus the cost and administrative burden of applying. This could diminish the effect of grant funding—grant-funded projects would contribute less to the firm's total R&D and firms may reallocate labor away from other grant-funded projects—but it also could put upward pressure on the estimates through a competition effect. Less constrained firms would be rationally more likely to substitute away from grants, as they have more resources to self-finance projects. If grants that would have gone to less constrained firms instead go to those that are more cost-sensitive, the average effect of grant funding may increase.

The first thing to note is that, all else equal, firms just *above* the grant rate threshold should be more likely to substitute away from grants, as the benefits of grant funding are lower. That said, without preferential treatment of small firms in award decision-making, changes in competition for funding due to some firms no longer applying could affect firms below the threshold as well. I do not have data on all applicants to explore this directly, but when including grant competition-level fixed effects in the estimation (Column 5 of Appendix Table C.24) the estimates increase, if anything.[61] This suggests that, if such substitution does occur and impacts the marginal effect of grant funding through a competition channel, it is putting downward pressure on the results as opposed to driving the large positive effects.

7.2.3 Cumulative Grant Funding

The marginal effect of grants also may increase over time due to cumulative funding. A grant might help expand previously-funded projects, or Innovate UK may preferentially treat previous winners to increase the likelihood of project success. However, when interacting the main treatment variables with the number of cumulative grants, I find that previous grant funding *dampens* the positive policy interaction effect (Column 6 of Appendix Table C.24).

[61] Industry-year fixed effects included in the baseline model also account for this behavior across competitions assuming that firms would be generally competing with others in their industry.

8 Conclusion

As a central driver of economic growth and competitiveness, innovation could play a pivotal role in reinvigorating economies following the productivity declines experienced in many countries since the 1970s. Policymakers are indeed revisiting their innovation policies and industrial strategies with support for R&D at the heart of it, but designing incentives so they deliver on their promise remains a long-standing challenge.

Many countries offer some combination of direct grants and fiscal incentives, and there is growing evidence that both subsidy types enhance R&D and innovation output. But as firms frequently tap into both and the choices organizations make in response to one instrument may depend on the availability of another, it is also important to understand their effects within the context of the broader innovation policy ecosystem. If they interact in their effects on firm behavior, policy reforms may augment or dampen the marginal returns to each subsidy type. This question has been under-explored so far.

In this paper, I present new evidence on whether grants and tax credits for R&D are complements or substitutes by implementing six quasi-experimental research designs that estimate the instruments' independent and interaction effects on R&D for firms in the UK. I find that they are complements for small firms, as increasing tax credit generosity greatly enhances the effect of grant funding, and financial constraints seem to be the key mechanism at play. The effects are much larger for constrained firms and the policy interactions increase both R&D "entry" as well as the propensity to invest for previous R&D performers. To corroborate the conclusion that financial constraints are a driving force behind these results, I also study the interaction effects for larger firms, which are less frequently constrained. I find that higher tax credit rates dampen the otherwise positive correlation between direct subsidies and R&D, suggesting that the instruments are substitutes for larger firms.

This paper also sheds light on the potential implications of policy interactions for the returns to business support programs and the importance of removing silos when designing policy more broadly. As policy interactions are ubiquitous across many economic settings, and as firm size-based policies are also common, the methods and results may also be of interest beyond the innovation context.

References

Agrawal, A., C. Rosell, and T. Simcoe, "Tax Credits and Small Firm R&D Spending," *American Economic Journal: Economic Policy*, 2020, *12* (2), 1–21.

Akcigit, U. and W.R. Kerr, "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

Alstadsaeter, A., S. Barrios, G. Nicodeme, A.M. Skonieczna, and A. Vezzani, "Patent Boxes Design, Patents Location, and Local R&D," *Economic Policy*, 2018, *33* (93), 131–77.

Arrow, K.J., "The Economic Implications of Learning by Doing," *Review of Economic Studies*, 1962, *29*, 155–173.

Ashraf, N., B.K. Jack, and E. Kamenica, "Information and subsidies: Complements or substitutes?," *Journal of Economic Behavior and Organization*, 2013, *88*, 133–139.

Azoulay, P., J.S. Graff Zivin, D. Li, and 2018 Sampat B.N., "Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules," *Review of Economic Studies*, 2018.

Bloom, N., J. Van Reenen, and H. Williams, "A Toolkit of Policies to Promote Innovation," *Journal of Economic Perspectives*, 2019, *33* (3), 163–184.

_ , R. Griffith, and J. Van Reenen, "Do R&D Tax Credits Work? Evidence from a Panel of Countries 1979-1997," *Journal of Public Economics*, 2002, *85* (1), 1–31.

Bronzini, R. and E. Iachini, "Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach," *American Economic Journal: Economic Policy*, 2014, *6* (4), 100–134.

Calonico, S., M. Cattaneo, and R. Titiunik, "Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Chen, Z., Z. Liu, J.C. Suarez Serrato, and 2021 Yi Xu D., "Notching R&D Investment with Corporate Income Tax Cuts in China," *American Economic Review*, 2021, *111* (7), 2065–2100.

Dechezleprêtre, A., E. Einiö, R. Martin, K-T. Nguyen, and J. Van Reenen, "Do Tax Incentives for Research Increase Firm Innovation? An RD Design for R&D, Patents, and Spillovers," *American Economic Journal: Economic Policy*, forthcoming.

Duflo, E., P. Dupas, and M. Kremer, "Education, HIV and early fertility: experimental evidence from Kenya," *Mimeo*, 2012.

Dupas, P., "What matters (and what does not) in households' decisions to invest in malaria prevention?," *American Economic Review: Papers and Proceedings*, 2009, *99* (2), 224–230.

Gaessler, F., B. Hall, and D. Harhoff, "Should there be lower taxes on patent income?," *Research Policy*, 2021, *50* (1).

Garicano, L., C. Lelarge, and J. Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, 2016, *106* (11), 3439–79.

Grembi, V., T. Nannicini, and U. Troiano, "Do Fiscal Rules Matter?," *American Economic Journal: Applied Economics*, 2016, *8* (3), 1–30.

Griffith, R., H. Miller, and M. O'Connell, "The UK will introduce a Patent Box, but to whose benefit? [Comment] IFS," 2010.

Guceri, I. and L. Liu, "Effectiveness of Fiscal Incentives for R&D: Quasi-experimental Evidence," *American Economic Journal: Economic Policy*, 2019, *11* (1), 266–91.

Hall, B.H. and J. Van Reenen, "How Effective are Fiscal Incentives for R&D? A Review of the Evidence," *Research Policy*, 2000, *29*, 449–69.

Hall, Bronwyn, "Tax Policy for Innovation," in Austan Goolsbee and Benjamin Jones, eds., *Innovation and Public Policy*, University of Chicago Press, 2022.

HMRC, "Research and Development Tax Credits Statistics," 2017. Available at: https://assets.publishing.service.gov.uk.

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

_ **, J. Rathje, J. Van Reenen, and J. Wong**, "Opening up Military Innovation: Causal Effects of Reforms to U.S. Defense Research," *NBER Working Paper 28700*, 2023.

Inderbitzin, L., S. Staubli, and J. Zweimuller, "Extended Unemployment Benefits and Early Retirement: Program Complementarity and Program Substitution," *American Economic Journal: Economic Policy*, 2016, *8* (1), 253–288.

InnovateUK, "Innovate UK funded projects since 2004 (data set)," 2023.

_ , "About Us," n.d.

Lach, S., "Do R&D Subsidies Stimulate or Displace Private R&D? Evidence from Israel," *Journal of Industrial Economics*, 2002, *50*, 369–390.

Mohnen, P., A. Vankan, and B. Verspagen, "Evaluating the Innovation Box Tax Policy Instrument in the Netherlands," *Oxford Review of Economic Policy*, 2017, *33* (1), 141–56.

_ **and L-H. Röller**, "Complementarities in innovation policy," *European Economic Review*, 2005, *49*, 1431–1450.

Nelson, R.R., "The Simple Economics of Basic Scientific Research," *Journal of Political Economy*, 1959, pp. 297–306.

OMB, "Analytic Perspectives: Budget of the U.S. Government, Fiscal Year 2024," 2023. U.S. Government Publishing Office, Office of Management and Budget.

ONS, "Comparison of ONS business enterprise research and development statistics with HMRC research and development tax credit statistics," 2022. Available at: https://www.ons.gov.uk/economy/governmentpublicsectorandtaxes/researchanddevelopmentexpenditur 09-291, Released 29 September 2022, Accessed August 2023.

_ , "Quality and Methodology Information Report, Business Enterprise Research and Development Survey QMI," 2022. Available at: https://www.ons.gov.uk/economy/governmentpublicsectorandtaxes/researchanddevelopmentexpenditur Accessed August 2023, ONS website last updated 17 November 2022.

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

S., Sorenson B.-Villegas-Sanchez C.-Volosovych V. Kalemli-Ozcan and S. Yesiltas, "How to Construct National Representative Firm Level Data from the ORBIS Global Dataset," *Working paper*, 2015.

UKRI, "General guidance," 2023. Available at https://www.ukri.org/councils/innovate-uk/guidance-for-applicants/general-guidance/contents-list, Accessed: 2023-09-01.

Zetlin-Jones, A. and A. Shourideh, "External financing and the role of financial frictions over the business cycle: Measurement and theory," *Journal of Monetary Economics*, 2017, *92*, 1–15.

Data References

Financial Analysis Made Easy (FAME), Bureau van Dijk. Available at: https://www.bvdinfo.com/en-gb/our-products/data/national/fame

Department for Business, Energy & Industrial Strategy, Innovate UK, UK Research and Innovation. (2017). Transparency Data: Innovate UK funded projects since 2004. Available at: https://www.gov.uk/government/publications/innovate-uk-funded-projects

Office for National Statistics. (2017). Business Expenditure on Research and Development, 1994-2015: Secure Access. [data collection]. 6th Edition. UK Data Service. SN: 6690, http://doi.org/10.5255/UKDA-SN-6690-6

Office for National Statistics. (2017). Business Structure Database, 1997-2017: Secure Access. [data collection]. 9th Edition. UK Data Service. SN: 6697, http://doi.org/10.5255/UKDA-SN-6697-9

MAIN TEXT TABLES

Table 1: R&D Tax Credit Rates for Small and Medium-Sized Enterprises

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: R&D Tax Relief Scheme Enhancement Rates and Benefits								
Year	Enhancement Rate	Payable Credit	Low Corp. Tax	Main Corp. Tax	Profit-Making % Benefit		Loss-Making % Benefit	Average % Benefit
					Low Tax	Main Tax	No Tax	
2005	0.50	0.16	0.19	0.30	0.095	0.150	0.240	0.152
2006	0.50	0.16	0.19	0.30	0.095	0.150	0.240	0.152
2007	0.50	0.16	0.20	0.30	0.100	0.150	0.240	0.154
2008	0.75	0.14	0.21	0.28	0.158	0.210	0.245	0.199
2009	0.75	0.14	0.21	0.28	0.158	0.210	0.245	0.199
2010	0.75	0.14	0.21	0.28	0.158	0.210	0.245	0.199
2011	1.00	0.13	0.20	0.26	0.200	0.260	0.250	0.235
2012	1.25	0.11	0.20	0.24	0.250	0.300	0.248	0.268
2013	1.25	0.11	0.20	0.23	0.250	0.288	0.248	0.263
2014	1.25	0.15	0.20	0.21	0.250	0.263	0.326	0.274
2015	1.30	0.15	0.20	0.20	0.260	0.260	0.334	0.278
2016	1.30	0.15	0.20	0.20	0.260	0.260	0.334	0.278
2017	1.30	0.15	0.19	0.19	0.247	0.247	0.334	0.269

Panel B: Averages and Changes Between "Low" and "High" Rate Periods

	Low TC Period Avg. (2005-2012)	High TC Period Avg. (2013-2017)	Percent. Point Difference	% Change
Enhancement Rate	0.750	1.280	0.530	0.707
Tax Credit Benefit	0.195	0.273	0.078	0.399

Notes: Panel A provides each component that goes into determining the R&D tax credit benefit for firms qualifying as small- and medium-sized enterprises (SMEs) under the R&D tax relief scheme receive as well as the calculated benefit rates. The inputs into the equation include R&D enhancement rates (Column 1), the payable credit rate for loss-making firms (Column 2), the corporate tax rate for profit-making firms making less than 300k in profits (Column 3), and the main corporate tax rate for firms making more than 300k in profits (Column 4). Columns 5 and 6 provide the resulting benefit rate for profit-making firms based on the lower or main corporate tax rate, Column 7 provides the benefit rate for loss-making firms, and Column 8 provides the average of all three benefit rates. Panel B provides the average enhancement and benefit rates for the low and high tax credit rate periods, the percentage point difference, and the percentage change.

Table 2: Innovate UK Grant Awards and Baseline Sample Summary Statistics

	Full Sample	**Baseline Sample**	**Wider Bandwidth**
	All Firms	MSE-Optimal Bandwidth	1 to 100 Empl.
	(1)	(2)	(3)
Panel A: Grant Awards			
No. of Unique Grants	8,230	561	2,006
No. of Unique Firms	5,301	417	1,498
Panel B: Funding Amounts			
Grant Amount (£000s)	£150.76	£242.99	£208.09
Total Project Cost Funded (%)	58.1%	53.0%	57.8%
No. of Observations	8,101	549	1,966
Panel C: Main Outcome Variable			
R&D Expenditures (£000s)	£42.00	£153.66	£120.69
No. of Observations	20,398	1,356	4,177

Notes: Table provides summary statistics for firms receiving Innovate UK grants between 2005 and 2017 conditional on satisfying the total assets and turnover criteria for being eligible for higher grant generosity rates as defined in Section 4.1. Samples include the year in which firms receive a grant and the three years that follow. Firms of all sizes (by employment) are included in Column 1. Column 2 restricts the sample to firms within the MSE-optimal window of 29 to 71 employees (in the year prior to winning a grant), which is the sample I use through my baseline regressions. Column 3 widens the bandwidth to include firms with 1 to 100 employees (in the year prior to winning a grant). To address outliers, I omit observations with the top 1% of non-zero R&D expenditures for firms with 1 to 100 employees in Columns 2 and 3.

Table 3: Independent Effect of Higher Grant Funding Rates on Small Firms' R&D

Dependent Variable:	R&D (1)	R&D (2)	R&D (3)	R&D (4)	R&D (5)
Panel A: Discontinuity in R&D Post-Grant					
1[Empl. < 50]	195.11*	184.45**	199.95**	219.54*	201.87*
	(109.49)	(91.20)	(99.95)	(126.20)	(117.41)
Observations	1,356	1,356	1,322	1,143	1,127
No. of Firms	417	417	407	375	373
Mean Dep. Var.	161.94	161.94	165.27	187.28	187.70
Panel B: No Discontinuity Pre-Grant					
1[Empl. < 50]	85.62	14.86	11.53	1.96	4.41
	(60.22)	(27.92)	(33.62)	(37.10)	(35.85)
Observations	1,281	1,281	1,258	1,093	1,089
No. of Firms	419	419	408	370	368
Mean Dep. Var.	72.61	72.61	73.86	63.46	63.69
R&D-Related Controls		x	x	x	x
Year FEs			x		
Industry FEs			x		
Industry-Year FEs				x	x
Additional Controls					x

Notes: Dependent variable is R&D expenditures (£000s). Table presents results from estimating independent direct effect of more generous grant funding rates for small firms on R&D over the full sample period (2005-2017). In Panel A, the sample includes "post-grant" observations (i.e., the years in which firms win grants and the three years that follow). In Panel B, observations for all years prior to when firms win their first grant are included (i.e., when there should be no difference at the 50 employee threshold). Baseline samples include firms within the MSE-optimal bandwidth (29 to 71 employees) using the year before they win a grant in Panel A and lagged employment in Panel B (conditional on meeting the total assets and turnover eligibility criteria for higher grant rates as well in both cases). All regressions include first-order polynomials of the running variable separately for each side of the threshold. Additional controls include lagged cumulative R&D expenditures, firm age, lagged cumulative R&D interacted with firm age, (log) total assets, and (log) current liabilities. In Panel A, the interaction and main effects of R&D prior to winning a grant interacted and the 50-employee threshold indicator are also included. Standard errors are clustered by firm. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table 4: Policy Interaction Effects on Small Firms' R&D

Dependent Variable:	R&D (1)	R&D (2)	R&D (3)	R&D (4)	R&D (5)
1[Empl. < 50] * Post 2012	310.79*	435.75*	415.51**	445.45**	560.28**
	(177.69)	(227.68)	(209.41)	(219.20)	(221.28)
1[Empl. < 50]	2.69	-194.98	-172.42	-188.27	-122.49
	(62.71)	(231.18)	(311.39)	(318.30)	(257.54)
Observations	1,356	1,291	1,291	1,269	1,047
No. of Firms	417	352	352	348	297
Mean Dep. Var.	161.94	157.59	157.59	158.63	186.70
Firm FEs		x	x	x	x
Year FEs		x	x	x	
R&D-Related Controls			x	x	x
Additional Controls				x	x
Year x Industry FEs					x

Notes: Dependent variable is total R&D expenditures (£000*s*). Sample includes observations for the grant year and the three years that follow for firms with 29 to 71 employees in the year before they win a grant (the MSE-optimal bandwidth around the threshold) and that also meet the total assets and turnover grant rate eligibility criteria. Running variable controls are included in all regressions. R&D-related controls include firms' average R&D prior to winning their first grant interacted with the 50-employee threshold as well as lagged cumulative R&D interacted with firm age along with each component individually. Additional controls include (logged) total assets and current liabilities (in real 2010 GBP). Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table 5: Larger Policy Interaction Effects for More Financially Constrained Firms

Dependent Variable:	R&D	R&D	R&D	R&D	R&D	R&D
Financial Constraint Proxy:	Short Term Debt		Operating Profit		Available Funds	
	Constrained	Unconstr.	Constrained	Unconstr.	Constrained	Unconstr.
	(1)	(2)	(3)	(4)	(5)	(6)
1[Empl. < 50] * Post 2012	787.90**	269.95	1402.15**	179.38	1407.33***	132.33
	(366.24)	(717.86)	(547.21)	(297.15)	(496.08)	(520.60)
1[Empl. < 50]	-1163.49*	675.75	-585.84	-199.17	-642.30	-819.12
	(653.98)	(585.16)	(568.70)	(335.11)	(505.40)	(578.14)
Observations	294	589	355	542	350	554
Mean Dep. Var.	193.55	204.15	379.40	74.40	335.41	106.04
Firm FEs	x	x	x	x	x	x
Year x Industry FEs	x	x	x	x	x	x
Baseline Controls	x	x	x	x	x	x

Notes: Dependent variable is R&D expenditures (£000*s*). Table provides results from estimating the subsidy interaction effects separately for firms under and over the median value of three financing constraint proxies (short-term debt, operating profit, and "available funds" (after-tax profits plus depreciation) in the year before winning a grant. More detail on these measures and the median values can be found in Appendix Table C.18. Effects for constrained firms—those under the median for operating profits and available funds and over the median for short-term debt—are in the odd-numbered columns and effects for less constrained firms are in even-numbered columns. Sample overall includes firms with 29 to 71 employees in the year prior to receiving a grant conditional on also meeting the turnover and total assets criteria for the higher grant rates. Running variable controls, fixed effects, and additional baseline controls are included in all regressions. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table 6: Policy Interaction Effects on R&D "Entry" and Propensity to Invest in R&D

Dependent Variable:	1[R&D>0]	R&D	1[R&D>0]	1[R&D>0]	1[R&D>0]
	(1)	(2)	(3)	(4)	(5)
Panel A: Effect of Policy Interactions					
1[Empl. < 50] * Post 2012	0.500***	1157.74***	0.176	0.293**	0.268***
	(0.154)	(370.14)	(0.136)	(0.114)	(0.092)
1[Empl. < 50]	-0.143	-233.59	-0.113	-0.220**	-0.224**
	(0.113)	(217.97)	(0.125)	(0.109)	(0.093)
Observations	711	711	1,047	1,610	2,428
Mean Dep. Var.	0.047	49.71	0.196	0.185	0.171
Panel B: Independent Effect of Higher Grant Funding Rates					
1[Empl. < 50]	0.032	198.404*	0.035	-0.013	0.006
	(0.043)	(108.747)	(0.065)	(0.051)	(0.043)
Observations	808	808	1,127	1,728	2,621
Mean Dep. Var.	0.062	73.961	0.189	0.179	0.164
Sample Empl. Range:	29 to 71	29 to 71	29 to 71	20 to 80	10 to 90
Sample Restriction:	No Prev. R&D	No Prev. R&D	None	None	None

Notes: Table presents results consistent with policy interactions increasing the probability of investing in R&D. Dependent variable in Column 1 and Columns 3-5 is an indicator equal to one if R&D expenditures are positive and zero otherwise, and in Column 2, it is R&D expenditures (*£000s*). In Columns 1 and 2, sample includes observations for which (lagged) cumulative R&D is zero. Otherwise, as in the baseline, sample includes observations for the year in which firms receive grants and the three years that follow conditional on meeting the turnover and total assets grant rate generosity eligibility criteria. The firm size ranges used for each regression are indicated in the bottom section of the table. All baseline running variable controls, fixed effects, and controls are included in all regressions. R&D-related controls are included when they are not absorbed by the FEs or do not drop out due to the sample restrictions. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table 7: Independent Effect of Tax Credit Policy on R&D Expenditures for Larger Firms

Dependent Variable:	R&D	R&D	R&D	R&D
	(1)	(2)	(3)	(4)
Estimation Approach:	**Regression Discontinuity**		**Diff-in-Disc**	
	Post-Policy	Pre-Policy		
1[Empl.<500]	1,000.01**	191.07		
	(446.19)	(154.52)		
1[Empl.<500] * Post 2008			539.48**	588.84**
			(238.93)	(241.67)
Observations	1,382	2,746	4,240	20,618
No. of Firms	584	1,299	1,061	5,493
Dep. Var. Mean	2,811	1,766	2,203	1,221
Sample Empl. Range	250-750	250-750	250-750	≤1000
Firm FEs			x	x
Year FEs			x	x

Notes: Table provides estimates for the effect of being eligible for more generous tax credit rates on larger firms' R&D expenditures (£000s). In Columns 1-2, I use an RDD to estimate the discontinuity in R&D in post-policy years (2009-2014) and pre-policy years (2000-2008). In Columns 3-4, I implement a difference-in-discontinuities design using data from all years. Sample includes firms with 250 to 750 employees in all regressions and the top 1% of the R&D distribution is omitted to account for outliers. Triangular weights are used in all cases and first-order polynomials of the running variable (employment) are included separately for each side of the threshold. The firm's average pre-policy R&D expenditures are included as a control in Columns 1, 3, and 4. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

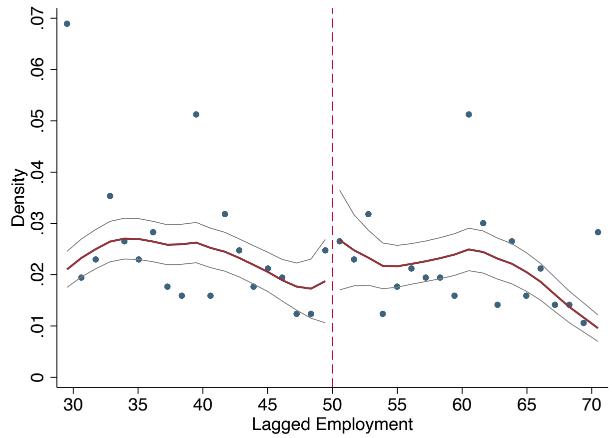
Table 8: Negative Policy Interaction Effects on Larger Firms' R&D Expenditures

Dependent Variable:	R&D (1)	R&D (2)	R&D (3)	R&D (4)	R&D (5)	R&D (6)	R&D (7)
Estimation Approach:			**Discontinuity-in-Effects**				**Diff-in-Disc-in-Effects**
Pre/Post Policy:	Post	Post	Post	Post	Pre	Pre	All Years
1[Empl.<500] * Subsidies	-4.54***	-2.83**	-2.93*	-3.92**	-0.38	-1.49	-0.20
	(1.15)	(1.42)	(1.60)	(1.75)	(2.32)	(1.44)	(0.88)
Direct Subsidies	5.95***	3.85***	3.74***	2.02**	3.23**	2.22***	2.17**
	(0.89)	(1.25)	(1.36)	(0.83)	(1.29)	(0.74)	(0.98)
1[Empl.<500]	1,264**	1,123**	934.0**	469.4	306.2	65.63	-305.52**
	(504.4)	(450.8)	(416.2)	(306.3)	(317.8)	(244.0)	(142.8)
1[Empl.<500] * Subsidies * Post 2008							-1.94**
							(0.85)
1[Empl.<500] * Post 2008							422.52**
							(186.00)
Observations	2,118	1,382	1,378	789	2,746	1,145	4,240
No. of Firms	1011	584	580	225	1299	346	1061
Dep. Var. Mean	2310	2811	2819	3653	1766	2400	2203
Pre-Policy R&D Control		x	x	x			
Industry FEs			x	x		x	
Year FEs			x	x		x	x
Firm FEs				x		x	x
Lagged R&D Control				x		x	x

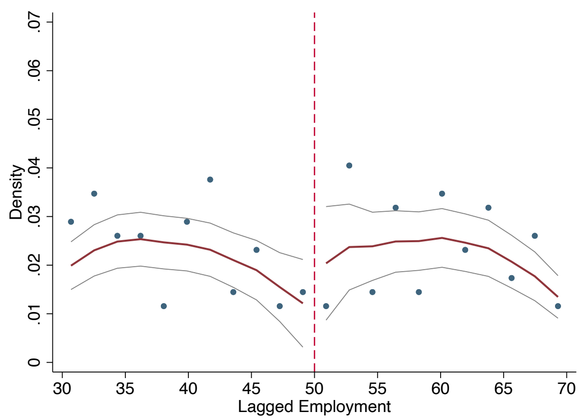
Notes: Table provides main results for the direct subsidies (£000s) and R&D tax credit interaction effects for larger firms around the 500-employee threshold. Sample includes firms with 250 to 750 employees in all regressions. In Columns 1-4, I estimate the discontinuity-in-effects model using data for the post-tax credit policy period (2009-14). The main regressor of interest is the interaction between direct subsidies and the indicator for whether the firm is under the tax credit generosity threshold. The pre-policy R&D control in Columns 2-4 is the firm's average pre-policy R&D expenditures. In Column 5, I use both pre- and post-policy data (2000-2014) and estimate the difference-in-discontinuities-in-effects by examining how the interaction effect changes between the pre- versus post-policy periods. The main regressor of interest in Column 5 is the three-way interaction. First-order polynomials of the running variable are included separately for each side of the threshold, the top 1% of the R&D distribution is excluded, and triangular weights are used in all regressions. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

MAIN TEXT FIGURES

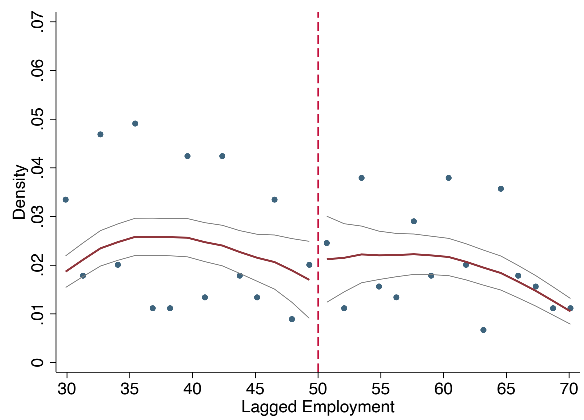
Figure 1: McCrary Tests for No Manipulation of Firm Size at the Small Firm Grant Generosity Threshold



(a) Full Sample Period (2005-2017)



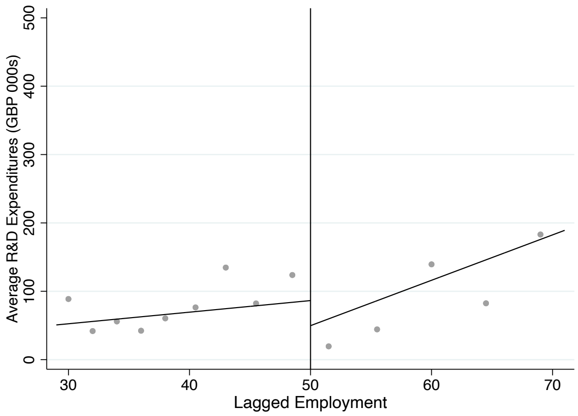
(b) Low Tax Credit Rate (2005-2012)



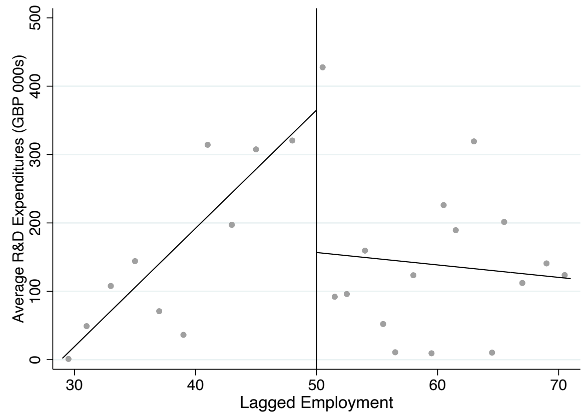
(c) High Tax Credit Rate (2013-2017)

Note: Figures provide McCrary tests for discontinuities in the distribution density of employment for the full sample period (Panel A) and then separately for the low tax credit rate period (Panel B) and high tax credit rate period (Panel C). Sample includes firms with 29 to 71 employees (lagged) that also meet the turnover and total assets grant generosity eligibility requirements. The discontinuity estimates (log differences in density height) at the small firm threshold (and standard errors) are 0.357 (0.335) in Panel A, 0.559 (0.623) in Panel B, and 0.300 (0.372) in Panel C. The discontinuities are not statistically different from zero and applying a simple t-test for statistical difference between the low and high tax credit period discontinuities produces a t-statistic of 0.357, suggesting that the diff-in-disc in the firm size distribution is also not statistically different from zero.

Figure 2: Discontinuities in Small Firms' R&D Expenditures at the Grant Rate Generosity Threshold Over Full Sample Period (2005-2017)



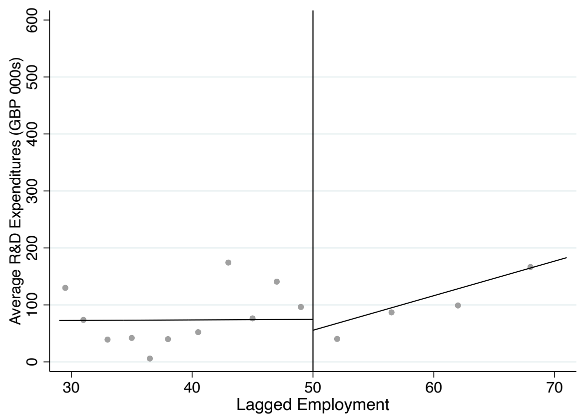
(a) Before Receiving Grants



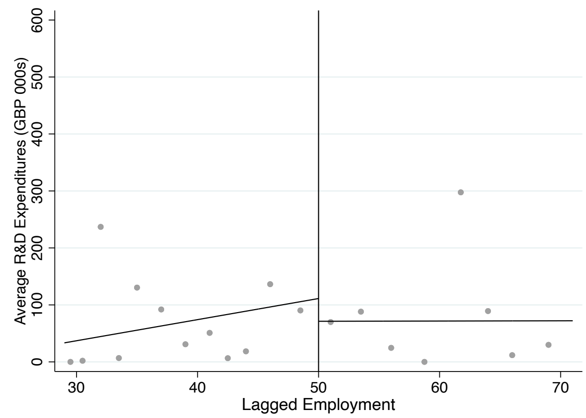
(b) After Receiving Grants

Note: Figure plots the discontinuities in R&D expenditures around the small firm size threshold of 50 employees before and after firms receive Innovate UK grants. Panel A includes observations for years before firms receive their first grant and Panel B includes observations for the year firms receive a grant and three years afterwards. Samples include firms with 29 to 71 employees (lagged) that also meet the turnover and total assets grant generosity eligibility requirements. Observations are binned following the IMSE-optimal quantile-spaced method. Local polynomials constructed using a triangular kernel, baseline control variables (besides pre-grant average R&D expenditures), and year fixed effects.

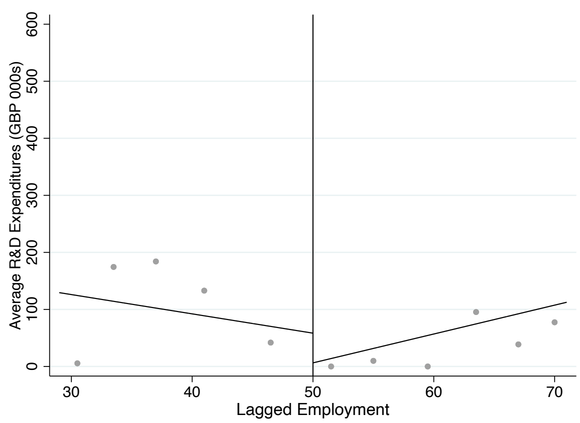
Figure 3: Policy Interaction Effects on Small Firms—Discontinuities in R&D at Grant Generosity Threshold Before and After Tax Credit Rate Increases



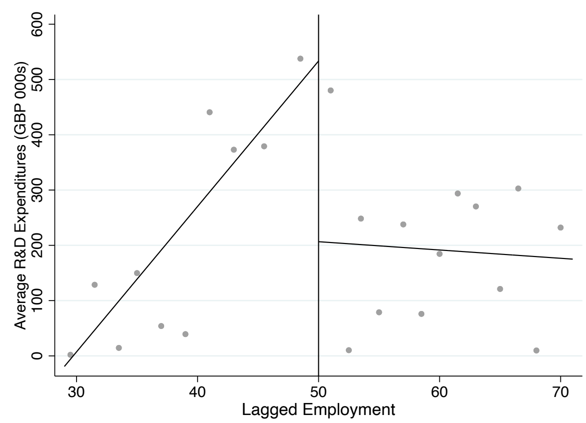
(a) Before Grant and Before TC Increase



(b) After Grant but Before TC Increase



(c) After TC Increase but Before Grant



(d) After TC Increase and After Grant

Note: Figure plots the discontinuities in R&D expenditures around the small firm threshold of 50 employees before after receiving Innovate UK grants and before and after the tax credit rate increases. Panel A includes observations for years before firms receive their first grant but before tax credit rates increase (2005-12). Panel B includes years when firms receive a grant and three years afterwards but before tax credit rates increase (2005-12). Panel C includes years after tax credit rates increase (2013-17) but before firms receive their first grant. Panel D includes years after tax credit rates increase (2013-17) and the year firms receive a grant as well as three years afterwards. Samples include firms with 29 to 71 employees (lagged) that also meet the turnover and total assets grant generosity eligibility requirements. Observations are binned following the IMSE-optimal quantile-spaced method. Local polynomials constructed using a triangular kernel, baseline control variables (besides pre-grant average R&D expenditures), and year fixed effects.

A Appendix: Data Preparation—For Online Publication Only

This appendix details the process I followed for preparing the data sets and the notable results associated with the matching.

A.1 Data Preparation for Small Firm Analysis

Direct Grants for R&D.—To study the effects of R&D grants and tax credits on small firms, I combine two main data sets. I focus on firms in the United Kingdom that received grants through Innovate UK, the UK's largest public funding body for private sector innovation. I start with the program's public database of all grants allocated since its inception and gather information on grants provided from 2005 through 2017.[62] The database contains details such as the grant award date, recipient, total grant award amount, proposed project costs, competition title and number, etc. Importantly, it includes unique company registration numbers (CRNs) so that firms can be matched over time and to other firm-level data sets.

I carry out a few data management and cleaning steps to prepare the data before matching to other data sets. I omit observations for which projects were withdrawn and observations without firm identifiers – which primarily include those associated with academic institutions, public service organizations (PSOs), charities, etc. and thus not entities to which the Innovate UK funding thresholds apply. Next, I search for and drop any remaining organizations of these types that did have some information in the firm identifier field, so they were not dropped in the first step, as well as grants provided to projects associated with the Catapult Network Centres, as these organizations tend to be also supported through a network that formally provides expertise and facilities to accelerate the application and scaling of their research. It is thus difficult to disentangle whether changes in R&D are associated with the grant funding itself or other resources from the Catapult Centres. I also omit grants noted as supporting activities like procurement and partnership building as well as others for which the grant funding thresholds do not apply, such as "vouchers," which are designed strictly for all smaller firms to seek expert advice.

The data set at this stage contains 16,250 observations across 7,492 organizations. I then examine the firm identifiers for obvious errors like non-standard formats and omit observations for which the CRN takes on the format of a non-UK country. I also drop observations that appear to include clear data entry errors in other grant-related variables,

[62] I use the "competition year" to determine the grant award date, and since these dates align with the application periods and span two calendar years, I use the second part of the competition time period, which is more likely to be associated with the year in which firms actually receive grants after the review process.

such as cases where the total amount of funding provided exceeds the project's proposed costs (as this would not align with the program's funding rules) and cases where the start of the project is later than one year after the grant year. This process drops only a small number of observations and leads to the data set containing 16,077 observations across 7,399 organizations.

Finally, since some organizations receive more than one grant per year, I collapse the data to the firm-year level for matching purposes, considering the "grant year" to be the second part of the competition's fiscal year. I sum quantitative variables such as grant amount and the number of grants received, and for non-quantitative variables, I use the value of the first observation for grants the firm receives that year, such as organization name and competition title.[63] The final clean Innovate UK data set that I match to other firm-level data contains 12,072 observations across 7,387 unique organizations for firms receiving grants between 2005 and 2017. On average, firms receive 1.63 grants over this time period.

Firm Balance Sheets and Profit and Loss Statement Data (Bureau van Dijk).—I collect firm-level financial data, such as R&D expenditures, employment, total assets, turnover, etc., from Bureau van Dijk's Financial Analysis Made Easy (FAME) Database, which is a commercial data set containing detailed balance sheet and P&L data on companies and unincorporated businesses in the UK and Ireland. Its construction begins with official filings content from the UK's Companies House and is enriched with additional efforts to ensure accuracy and fill in some gaps. Overall, the database covers over 11 million companies, including 2 million in a detailed format, 1.3 million companies that are active but have not yet filed accounts or are not required to file, and 6 million companies that are no longer active.

In the UK, all limited, PLC, LLP and LP companies are required to file accounts with Companies House, so the data capture at least basic information on the universe of firms in the UK, representing about 1 million companies as of 2015 in FAME (Kalemli-Ozcan and Yesiltas 2015). All companies in the UK must keep accounting records of all money received and expended, assets, and liabilities and report this information. However, as I discuss more below, small firms had the option to report less detailed accounts over my study period. The FAME data provides the latest account date, but some firms report quarterly. I follow Kalemli-Ozcan and Yesiltas (2015) and define the filing year based on the year of the latest filing date if the date is June 1 or later, and otherwise, I use the preceding year.

Matching Innovate UK Grant Data to Bureau van Dijk's Financial Data.—I next match the prepared firm-year Innovate UK grant data with the balance sheet and P&L statement data

[63]This limits the usefulness of some of the non-quantitative grant-specific variables, so I primarily only use it when carrying out a few robustness checks, like when including competition-level fixed effects.

from FAME. Of the 12,072 firm-year observations in the grant data, 11,750 observations (97.3%) across 7,146 organizations match when merging on company registration numbers. Notably, the number of observations that did not match in years that I define as being the "low tax credit rate" (2005-2012) and "high tax credit rate" (2013-2017) periods when studying policy interactions for small firms is very similar.[64]

I apply a few standard data preparation rules to be sure that I omit potential data entry errors for key financial variables. For example, I drop observations for which variables like total assets are negative or when the observed year is earlier the firm's founding year, and I convert monetary figures into real 2010 terms using the World Bank's Consumer Price Index. Once this is complete, I again drop any observations associated with firms that did not receive a grant from 2005 to 2017 in this remaining "clean" sample. The data set at this stage is an unbalanced panel of 72,481 observations across 7,035 firms that received Innovate UK grants between 2005 and 2017, and these firms received an average of 1.72 grants over this time period.

Throughout the majority of the analyses when examining the effects of grant funding and policy interactions on small firms, I limit the sample to just the year in which firms receive a grant and the three years that follow. This provides a data set of 28,796 observations across 7,035 firms. Furthermore, for most of the estimations throughout the paper, I limit the sample to firms that also meet the total assets and turnover criteria for higher grant rate generosity as described below, which reduces the sample to 20,398 observations across 5,301 firms. The sample size then declines more substantially when restricting the data to firms with non-missing employment data in the year prior to receiving a grant, which is required for defining treatment status. I discuss this and the implications below as well as in the main text.

Advantages of Using FAME for Small Firms.—There are two key advantages of using FAME to study small firms relative to some other R&D data sets frequently used to study innovation in the UK that are important for my paper, such as the Business Enterprise Research and Development (BERD) survey. First, BERD has been shown to vastly under-cover small firms, and second, the sampling population conditions on already being an R&D performer. I discuss these challenges more below.

One of the key advantages of using FAME to study small firms over BERD is that BvD starts with a comprehensive list of firms from Companies House (and the data reported therein) and then enhances the officially-reported data with its own research. On the other hand, while BERD is well-suited for studying larger firms—which has been used in many

[64] Only 189 observations did not match in the "low" period and 151 did not match in the "high" period.

other studies of UK innovation policies and which I indeed use to study larger firms as well—and especially those that are already R&D performers, small firms are vastly under-covered. The sampling frame that ONS uses to identify firms for the BERD survey is a list of all known R&D performers in Great Britain, whereby R&D-performers are identified from responses on other business surveys (like the Annual Business Survey (ABS)). The top 400 businesses according to the size of previously reported R&D expenditures are selected automatically and then another 3,600 firms are randomly selected from the remaining firms on the pre-determined list of R&D performers (ONS 2022b).

Importantly, most of these "feeder surveys" from which firms are sampled carry out a census of all large businesses (i.e., those with more than 250 employees by their standard definition), so even though the random sample may consist of firms with lower R&D expenditures than the "biggest" R&D-performers, many small firms according to standard definitions are never included in these census surveys and thus are never covered by the BERD survey. Instead, data for non-sampled businesses are estimated and imputed using a ratio method based on employment.

The under-coverage of small firms in BERD is striking. The HMRC's official R&D statistics historically indicated that R&D is higher than what BERD statistics suggest. For example, for the financial year 2020/21, R&D expenditures were 42% higher than the BERD estimates. To try to better-understand this divergence, the ONS and HMRC carried out a microdata sharing project and concluded that the overwhelming majority of it could be explained by the under-coverage of small firms (ONS 2022a). Using their newly-developed "uplift" method to assess the degree of under-representation, the report finds that the value of expenditures performed by UK firms would have been 43 billion GBP compared to the 26.9 billion GBP previously reported if the necessary methodological adjustments were made. Likewise, R&D expenditures would have been 15 and 15.6 billion higher in 2018 and 2019. While there are some other differences between HMRC's and BERD's data collection and R&D measurement methods that can explain small proportions of this difference, the report concludes that most of the difference is due to the small firm under-coverage issue.

Second, the sampling approach used to construct the BERD survey population conditions on firms already being R&D performers. Part of the objective of grant programs is to help firms overcome financial constraints, which are more likely to be firms with no previous R&D, and thus many firms of interest were likely never included in BERD. Part of my objective also is to study whether the combination of policies induced R&D entry—an indication of potentially overcoming financial constraints—which could not be done if all firms in the sample were already investing in R&D.

Final variable construction.—Lastly, I construct a few final variables required for the analyses. I convert monetary variables into real 2010 terms using the World Bank's CPI for GBP, and I convert total assets and turnover into euros using each year's average exchange rate for the purposes of defining treatment status, as the Innovate UK generosity thresholds are defined using euros. For R&D expenditures, I consider all missing data to be zeros. I discuss the rationale and implications for this more below and in the main text and I also carry out a number of robustness checks to help ensure this assumption is not biasing the results. Essentially, though, the idea is that because I am already conditioning on firms that choose to report employment (since employment is the running variable), missing data for financial variables likely represent true zeros, as these pieces of information likely would have been reported as well if firms already choose to report employment.

To define whether the firm is treated and benefits from more generous grant rates, I use the values of employment, total assets, and turnover from the year prior to receiving a grant. To be eligible for the higher grant rates, firms must have fewer than 50 employees and either total assets or turnover must be lower than 10m euros. Since employment is the binding criteria, I restrict the estimation sample to include only firms that meet the total assets and/or turnover requirements in the year prior to receiving a grant and then define treatment using the 50-employee threshold. This allows me to take all three eligibility criteria into account and to rely on the binding criteria for treatment status. If both turnover and total assets data are missing in the year prior to receiving a grant, I keep them in the data set to preserve sample size but show that the results are robust to not imposing the turnover and assets restrictions as well.

I apply the same treatment status for the three years that follow grant receipt when constructing the treatment and running variables. That is, if a firm receives a grant in 2014, I use the 2013 employment, total assets, and turnover values to define treatment, and if it meets the eligibility requirements for the small firm grant generosity rate in 2013, I consider the firm treated in 2014, 2015, and 2016 as well (for both the treatment indicator and running variable construction). If the firm receives another grant during those years, I update the treatment status using the same procedure, replacing the original treatment variables with updated ones reflecting status associated with the most recent grant received.

Variable coverage.—A 97.2% match rate between the Innovate UK and FAME data suggests that BvD's coverage of firms in FAME is quite comprehensive, even for relatively small firms. This is because BvD starts by drawing company information from Companies House, where all firms in the UK are required to file accounts with at least some basic information. At the same time, this matching rate does not reflect the coverage in my data set, since not all variables are fully populated, even with BvD's additional research supplementing the Companies House data. This is because not all firms are required to report all information.

The first variable for which this is particularly important is employment, the running variable used in the research designs described in Section 3. There is indeed a significant portion of observations for which employment data are missing. Of the 28,796 observations across 7,035 firms in the starting "cleaned" sample (which includes years in which firms receive grants and the three years that follow), current employment information is included for 15,111 observations (52.5%) across 4,756 (67.6%) of firms. Lagged employment is populated for 13,037 observations (45.3%) across 4,204 firms (59.8%). When considering lagged employment in the year prior to receiving a grant, which is what I use to define treatment status and then apply to all years associated with the grant and is thus most relevant for my estimation, 11,482 observations (40%) across 2,975 firms (42.3%) have populated data.

The most relevant sub-sample to consider in terms of coverage rates is the final starting data set containing 20,398 observations across 5,301 firms that I use throughout most of the paper (conditioning on also meeting the total assets and turnover eligibility requirements for more generous grant rates, as described above). Of this data, 4,703 observations (23%) across 1,614 firms (30.4%) have employment data for defining treatment status. Although these coverage rates might seem low, the sample size for matched firms with employment information is still much larger than when attempting to match Innovate UK data to BERD data that is not imputed.[65] It is also likely a more representative sample of the types of firms I aim to study since it does not condition on already being an R&D performer.

The main potential concern that missing data introduces from an identification strategy perspective is that firms under the 50-employee threshold that choose to report certain information may differ systematically relative to those that do not. As such, the second variable for which it is important to consider missing data is R&D, the main outcome of interest. Throughout my analyses, I assume all missing R&D data are zeros. See the main text of the paper for a discussion of what this implies for my estimation. The main potential concern with doing this is that small firms in the UK can opt out of reporting all information on P&L statements. It is first worth noting that firms in my baseline estimation sample are more likely to report R&D (if they do invest) because I already inherently condition on reporting employment given the treatment variable definition. If firms report employment, missing information for R&D is more likely to represent a true zero. That said, if firms under the 50-employee threshold do choose to not report R&D even though they report employment, this could put either upward or downward pressure on the estimates if they systematically differ from firms that do not report R&D (yet do invest). I carry out tests that suggest it is likely not a threat in my setting in Section 5.4 and Section 6.3.

[65] I discuss the limitations of using BERD to study small firms in more detail above.

A.2 Data Preparation for Larger Firm Analysis

UK Data Services Secure Lab.—The regression analysis for large firms entails linking several microbusiness datasets that are legally protected and held by the UK's Office of National Statistics (ONS). Accessing the data requires a special process, which begins with training and taking an exam regarding the use and protection of sensitive data to become a UK Accredited Researcher. A research proposal then must be submitted and approved, justifying the use of the data sets and providing the reasons that they must be accessed and linked in order to answer a question that is relevant for the UK's public good. Once approved, all data use and analysis must be conducted in the UK Data Service Secure Lab.

Firm R&D Expenditures.—I use data from Business Enterprise Research and Development (BERD) survey to study R&D expenditures of large firms. The BERD survey is conducted by the ONS following the Frascati Manual methodology (OECD 2002), collecting information on R&D expenditures and other characteristics of firms identified as actively performing R&D. A stratified sampling approach is employed to select which enterprises will receive a BERD questionnaire. The ONS primarily uses the Annual Business Survey (ABS) to identify R&D-performing firms as well some other data sources such as the UK Community Innovation Survey and HMRC data on firms claiming R&D tax credits.

All questionnaires sent to those selected include a minimum set of questions on total R&D spending and R&D employment. The largest spenders on R&D receive "long form" questionnaires and the remainder receive a "short form." The short form asks for basic information related to R&D, such as in-house and extramural expenditures and total headcount of R&D employees. The long form covers more detailed information, such as how R&D expenditures are spent based upon capital and non-capital expenditures. Enterprises not included in the stratified sampling, and responses to questions on the long form from firms that were just sent a short form, have imputed values. These are the mean values of the variable as a share of employment in the firm's size band-sector group.

I collect BERD data from 2000 through 2014 and omit defense-related R&D investments. The full BERD datasets begin with about 30,000 observations per year. I take a number of steps to prepare the data for analysis. First, I do not use imputed values in order to avoid introducing measurement error. Omitting observations with imputed responses for the key outcome variable of interest (R&D expenditures) reduces the sample size to about 2,500 observations per year. Next, I omit observations where the Inter-Departmental Business Registrar (IDBR) reporting unit number seems as though it was recorded incorrectly due to taking on the wrong format. I also drop observations where the IDBR is duplicated, as there is no consistent way of understanding which entry is correct when the responses do not align. In total, this process results in dropping <0.01 percent of the observations.

Finally, the BERD responses are observed at the IDBR reporting-unit level, but funding and tax credit eligibility rules are determined by firm characteristics at the "enterprise group" level, which is a larger statistical unit. The EU Regulation on Statistical Units defines enterprise groups as "an association of enterprises bound together by legal and/or financial links" (EEC 696/93). The reporting unit level is associated with a geographical unit, whereas enterprise groups capture all reporting units associated with an enterprise.

The BERD datasets for each year include all reporting unit-year observations that were identified by ONS as firms performing R&D in the UK, yet the assignment to treatment in this analysis depends on whether the enterprise group satisfies the eligibility criteria. I aggregate the BERD data to the enterprise group level so that it can be matched to the Business Structure Database (BSD), which provides data on the enterprise group's total employment, and so that the R&D expenditure data captures the entire enterprise group's R&D investment levels. Furthermore, the location where R&D funds are allocated to an enterprise might not be the same local-level reporting level that is observed in BERD. This aggregation process results in only a very small further reduction in the sample size (usually less than 100 observations per year). The final BERD data set consists of about 2,000 to 2,500 enterprise groups per year from 2000 through 2014.

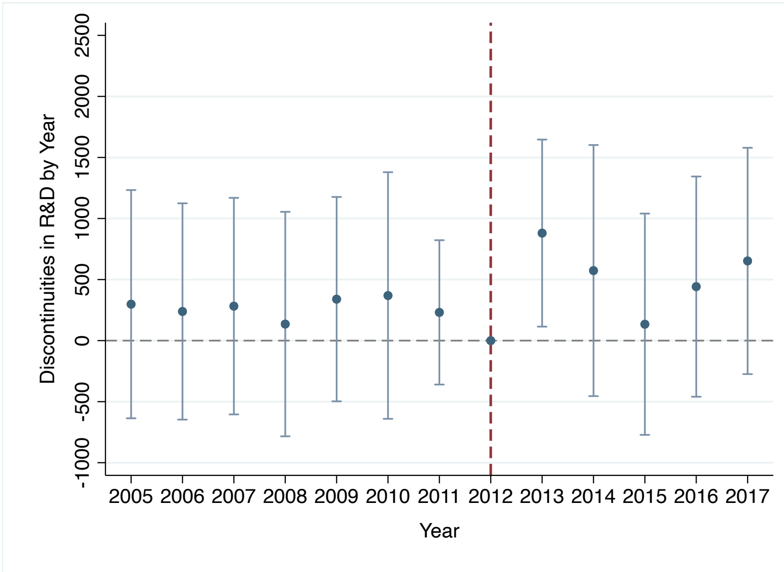
Determining Funding Level Eligibility.—I use the UK's Business Structure Database (BSD) to determine each enterprise group's tax credit rate eligibility. The BSD is also held securely by the ONS and requires UK Data Services Secure Lab access. It includes information on a small set of variables for nearly all businesses in the UK, and since it allows for one to observe a reporting unit's enterprise group, I use this to determine each enterprise group's employment level and thus tax credit rate eligibility. The data are derived mostly from the IDBR, which is a live register of administrative data collected by HM Revenue and Customs including all businesses that are liable for VAT and/or has at least one member of staff registered for the Pay As You Earn (PAYE) tax collection system. The BSD only misses very small businesses, such as those that are self-employed, and covers almost 99 percent of the UK's economic activity.

The BSD annual datasets include variables such as local unit-level and enterprise-level employment, turnover, company start-up date, postcodes, and the Standard Industrial Classification (SIC). I aggregate variables to the enterprise group level. If the observation is missing an enterprise number and does not belong to a larger enterprise group, I use the given observation's values for each variable. There are about 3 million observations per year. The enterprise group numbers are anonymous but unique so that they can be linked to other data sets held by the ONS.

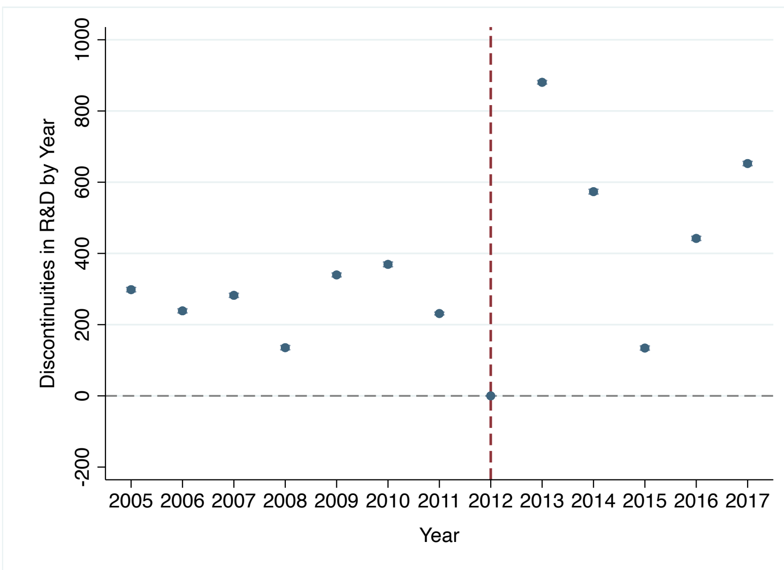
Final Data Sample Preparation.—A few final steps are taken to prepare the data. First, all expenditure and financial variables are converted into real 2010 terms using the World Bank's Consumer Price Index. Observations associated with inactive firms are dropped from the sample, which results in dropping only 72 observations, and I omit the top 1% of the R&D distribution to address the highly skewed nature of R&D investments. The final data set includes about 2,000 to 2,500 firms/enterprise groups per year from 2000 through 2014.

B Appendix: Additional Figures—Online Publication Only

Figure B.1: Event Study Plot of Discontinuities in R&D by Year



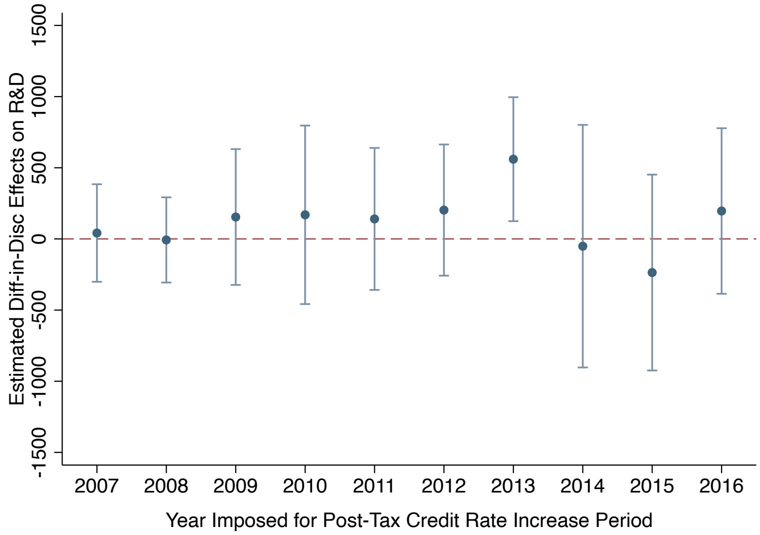
(a) **Estimates with 95% Confidence Intervals**



(b) **Estimates without Confidence Intervals**

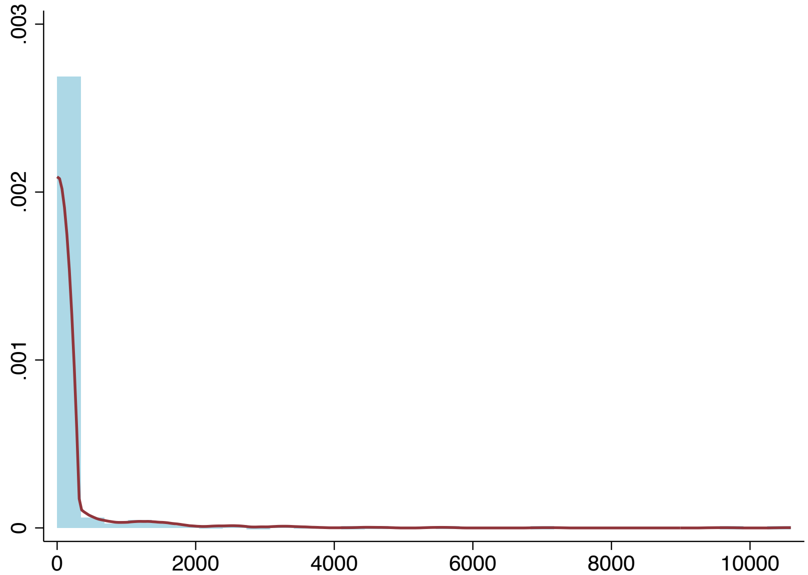
Note: Figure plots yearly discontinuities in R&D, which are found by estimating the diff-in-disc model of Equation 2 but interacting the treatment dummy and running variables with indicators equal to one for each year (using 2012 as the reference period). All fixed effects and controls are included in the regression. Panel A plots the results with 95% confidence intervals and Panel B omits the confidence intervals so the difference in coefficient estimate magnitudes is more transparent. Coefficient estimates and their standard errors are provided in Appendix Table C.9 for reference. The average of the point estimates in the low tax credit rate period is 270.7 and the average in the high tax credit rate period is 536.6.

Figure B.2: Placebo Tests for Tax Credit Rate Increase Timing



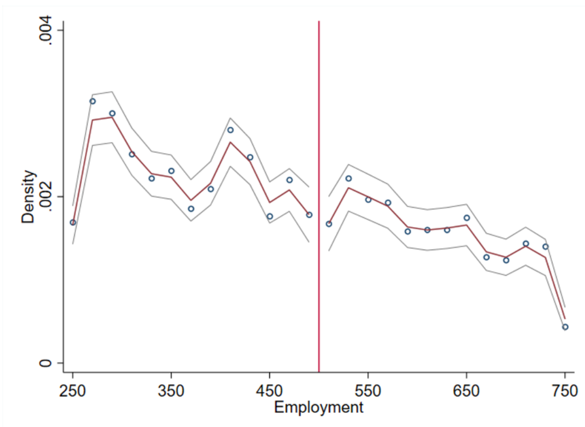
Note: Figure plots diff-in-disc estimates and their 95% confidence intervals from when estimating separate equations that impose different years as "pseudo" policy change years. For example, the 2007 coefficient is associated with estimating the diff-in-disc model of Equation 2 but assuming that tax credit rates enter a high tax credit rate period in 2007 rather than 2013. As expected, the only case in which there is a large, positive, and statistically significant diff-in-disc is for the actual treatment year (2013).

Figure B.3: Distribution of R&D Expenditures for Baseline Sample

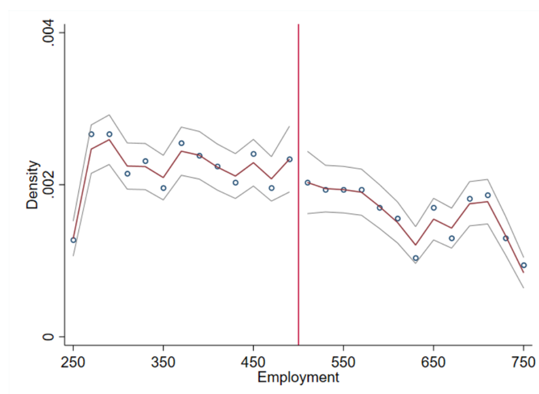


Note: Figure plots R&D expenditure distribution and kernel sensity function (epanechnikov) for baseline estimation sample of firms with 29 to 71 employees including observations in the year in which firms receive grants and the three years that follow and after omitting the top 1% of the non-zero R&D distribution.

Figure B.4: McCrary Tests for No Manipulation of Firm Size at the "Larger" Firm Tax Credit Rate Generosity Threshold



(a) Pre-Policy Period (2000-2007)



(b) Post-Policy Period (2009-2014)

Note: Figures provide McCrary tests for discontinuities in the distribution density of employment for firms with 250 to 750 employees in the pre-tax credit rate policy period (2000-2007) in Panel A and post-tax credit rate policy period (2009-2014) in Panel B. The discontinuity estimates (log differences in density height) at the 500-employee firm threshold (and standard errors) are -0.063 (0.284) in Panel A and -0.141 (0.289) in Panel B. The discontinuities and the difference in discontinuities are not statistically different from zero. Figure was created using data from the UK's Business Structure Database and Business Enterprise Research and Development Database, Department for Business, Innovation, and Skills, Office for National Statistics.

C Appendix: Additional Tables—Online Publication Only

Table C.1: R&D Tax Credit Rates for "Large" Firms Over the Tax Credit Generosity Threshold of 500 Employees Relative to Under

	(1)	(2)	(3)	(4)	(5)
				Difference from SMEs	
Year	**Enhancement Rate Rate**	**Corporate Tax**	**Tax Credit Benefit**	**Pct. Point in Enh. Rates**	**Pct. Point in Benefits**
2008	0.3	0.28	0.084	-0.45	-0.115
2009	0.3	0.28	0.084	-0.45	-0.115
2010	0.3	0.28	0.084	-0.45	-0.115
2011	0.3	0.26	0.078	-0.70	-0.157
2012	0.3	0.24	0.072	-0.95	-0.196
2013	0.3	0.23	0.069	-0.95	-0.194
2014	0.3	0.21	0.063	-0.95	-0.211
Averages	0.3	0.254	0.076	-0.700	-0.158

Notes: Table provides R&D Tax Relief Enhancement Rates for firms with more than 500 employees (Column 1), the UK's main corporate tax rate (Column 2), the R&D tax credit benefit based on the policy's formula using enhancement rates and corporate tax rates (Column 3), and the percentage point differences for these "large" firms and SMEs (just under the threshold) in enhancement rates (Column 4) and tax credit benefit (Column 5).

Table C.2: Descriptive Statistics of Data Used for Larger Firm Analyses

Firms in Sample:	**All Firms**		**Below Threshold**		**Above Threshold**	
Pre/Post Tax Credit Rate Policy	Pre	Post	Pre	Post	Pre	Post
	(1)	(2)	(3)	(4)	(5)	(6)
Mean R&D expenditures (£000s)	1751.3	2288.7	1815.2	2353.0	1656.4	2200.4
	(3436.2)	(4492.8)	(3540.8)	(4782.6)	(3274.0)	(4061.8)
Mean R&D direct subsidies (£000s)	22.49	61.98	22.94	69.62	21.82	51.48
	(90.37)	(322.68)	(96.75)	(394.38)	(79.98)	(192.64)
Mean prop. of R&D subsidized (%)	0.030	0.046	0.031	0.046	0.029	0.045
	(0.084)	(0.086)	(0.084)	(0.089)	(0.084)	(0.082)
Observations	2746	2118	1641	1226	1105	892
No. of Firms	1299	1011	900	702	573	445

Notes: Table provides descriptive statistics of data used to study larger firms for firms with 250 to 750 employees from the UK's Business Enterprise Research and Development (BERD) survey and Business Structure Database (BSD). The full baseline sample is included in Columns 1-2, firms below the R&D tax credit generosity threshold of 500 employees are included in Columns 3-4, and firms above the threshold are included in Columns 5-6. Sample omits top 1% of the R&D expenditure distribution. Pre-policy period statistics (2000-2008) are provided in the odd-numbered columns and post-policy period statistics (2009-2014) are in even-numbered columns. Standard errors are in parentheses.

Table C.3: Covariate Balance Around Grant Rate Threshold Before Receiving Grants

Dependent Variable:	R&D	Cumulative R&D	Age	log(Total Assets)	log(Current Liabilities)
	(1)	(2)	(3)	(4)	(5)
Panel A: Discontinuities in Pre-Grant Years					
1[Empl. < 50]	85.621	232.945	2.517	0.001	0.201*
	(60.217)	(172.179)	(2.588)	(0.105)	(0.121)
Observations	1,281	1,281	1,281	1,281	1,279
Mean Dep. Var.	72.610	163.468	21.564	8.390	7.576
Panel B: Differences-in-Discontinuities in Pre-Grant Years					
1[Empl. < 50] * Post 2012	-0.549	-114.409	8.262	-0.226	-0.690
	(79.756)	(407.362)	(6.645)	(0.344)	(0.492)
1[Empl. < 50]	86.509	254.292	1.654	0.025	0.278**
	(66.664)	(191.197)	(2.891)	(0.118)	(0.139)
Observations	1,281	1,281	1,281	1,281	1,279
Mean Dep. Var.	72.610	163.468	21.564	8.390	7.576

Notes: Table provides discontinuities (Panel A) and differences-in-discontinuities (Panel B) in covariates when using observations from only the years before firms receive their first Innovate UK grant. Coefficients are from estimating Equations 1 (Panel A) and 2 (Panel B) with different covariates as dependent variables conditional on first order polynomials of the running variable (included separately for each side of the threshold). The dependent variable is R&D in Column 1, lagged cumulative R&D in Column 2, firm age in Column 3, (log) total assets in Column 4, and (log) current liabilities in Column 5. Sample includes firms within the MSE-optimal window of 29 to 71 employees that also meet the turnover and total assets grant rate generosity criteria. Standard errors are clustered by firm. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.4: No Discontinuities in Likelihood of Receiving a Grant

Dep. Var.:	Received (1)	Received (2)	Received (3)	Received (4)	Received (5)	Received (6)
1[Empl. < 50]	-0.020	-0.018	-0.027	-0.021	-0.016	0.048
	(0.021)	(0.024)	(0.026)	(0.026)	(0.029)	(0.038)
1[Empl. < 50] * Post 2012				-0.008	-0.019	-0.079
				(0.043)	(0.049)	(0.070)
Observations	4,510	3,554	3,268	4,510	3,554	2,954
Mean Dep. Var.	0.149	0.154	0.157	0.149	0.154	0.159
Cond. on All Eligibility Criteria		x	x		x	x
Baseline Controls and FEs			x			x

Notes: Dependent variable is an indicator equal to one if the firm received a grant that year and zero otherwise. Data for all years (2005-2017) included for firms with (lagged) employment of 29 to 71 employees conditional on also meeting the total assets and turnover grant rate generosity criteria. In Columns 1 and 4, the sample is not yet limited to firms that also meet the turnover and total assets grant rate generosity criteria. In Columns 2, 3, 5, and 6, the other eligibility criteria are applied. Running variable controls are included in all regressions and the remaining controls and fixed effects of the baseline specification are included in Columns 3 and 6. Standard errors are clustered at the firm level. Asterisks denote **p* <0.10, ***p* <0.05, ****p* <0.01.

Table C.5: Sample of UK Policies Providing Benefits for Smaller Firms

Policy/Program	Description	Firms Affected
Small Business Rate Relief	Relief from property business rates charged on non-domestic properties.	Firms with rateable value less than £15k or business uses only one property.
Corporate Taxes	Single Corporation Tax rate of 20% for non-ring fence profits.	Determined by profits as opposed to turnover, employment, or total assets.
Employment Allowance	Discount on National Insurance bill.	Any business paying employers' Class 1 National Insurance
Venture Capital Schemes: Enterprise Investment Scheme, Seed Enterprise Investment Scheme, and Social Investment Tax Relief	Tax relief provided to investors of venture capital schemes. Relief provided against income tax or capital gains tax.	Tax relief is provided to investors as opposed to firms.
Enterprise Investment Scheme	A venture capital scheme that helps companies raise money.	Firms must not have gross assets worth more than £15m before shares are issued (and £16m afterwards), and must have fewer than 250 employees.
Seed Enterprise Investment Scheme	A venture capital scheme that helps companies raise money.	Firms must not have gross assets worth more than £200k at the time when shares are issued, and must have fewer than 25 employees.
Small Business: GREAT Ambition	Commitment to help small businesses grow.	No firm size definitions that align with the Innovate UK definitions.
British Business Bank	A business development bank committed to making finance markets work better for small businesses.	Support programs for start-ups and small businesses in general with no noticeable advantages to firms that align with the firm size definitions for grant generosity.
Employer NI Contributions	Employers pay secondary national insurance contributions to HMRC.	Rates are determined by profits as opposed to employment, turnover, or total assets.
Value Added Tax	VAT registration is required for firms of a certain size.	The threshold for VAT registration is £85k.
Pay As You Earn	Payment by employers as part of the payroll so that the HMRC can collect income tax and national insurance.	Income tax rates depend on how much of taxable income is above personal allowance, and rates are determined by earnings.
Export Credits Guarantee Scheme	Encourages exports by SMEs by ensuring successful implementation of scheme.	Applies to all SMEs, not just small firms.
Loan Guarantees for SMEs	Government agreement with large banks to extend loans to small businesses in the UK, increasing the availability of finance.	Applies to all SMEs, not just small firms.
Enterprise Capital Funds	Financial schemes to address the provision of equity finance to certain firms and to invest in high growth businesses.	Applies to all SMEs, not just small firms.
Business Angel Co-Investment Fund	A £100M investment fund for UK businesses.	Applies to all SMEs, not just small firms.

Notes: Table provides information on a sample of other policies in the UK that provide incentives for small businesses. No policies that could confound the diff-in-disc estiamtes for small firms are found.

Table C.6: No Discontinuities in Likelihood of Reporting Profits or R&D Before Receiving First Grant

Dep. Var. (indicator):	Reported (1)	Reported (2)	Reported (3)	Reported (4)	Reported (5)	Reported (6)
1[Empl. < 50]	-0.020	-0.032	-0.045	0.047	0.036	-0.001
	(0.023)	(0.028)	(0.029)	(0.039)	(0.047)	(0.042)
Observations	1,623	1,281	1,089	1,623	1,281	1,089
Mean Dep. Var.	0.957	0.947	0.953	0.093	0.092	0.100
DV = 1 if Reported Profits	x	x	x			
DV = 1 if Reported R&D				x	x	x
Cond. on All Eligibility Criteria		x	x		x	x
Baseline Controls and FEs			x			x

Notes: Dependent variable is an indicator equal to one if there is non-missing data for the (before tax) profits or R&D variables. Missing data could reflect either firms not reporting that information in the P&L statements that they file or BvD not gathering the additional information for that firm. Sample includes only years prior to when firms receive their first grant. In Columns 1 and 4, all firms in the baseline sample with (lagged) employment of 29 to 71 but without conditioning on the grant generosity threshold turnover and total assets criteria also being met. In Columns 2, 3, 5, and 6, the other eligibility criteria are applied. Running variable controls are included in all regressions and the remaining controls and fixed effects of the baseline specification are included in Columns 3 and 6. Standard errors are clustered by firm. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.7: Pseudo-Threshold Tests for RDD Grant Funding Rate Effects

Dependent Variable:	R&D (1)	R&D (2)	R&D (3)	R&D (4)	R&D (5)
Panel A: Triangular Weights and Even Bandwidth					
Pseudo-Thresh (5)	-14.94				
	(27.88)				
Pseudo-Thresh (10)		-60.66			
		(43.21)			
Pseudo-Thresh (15)			-19.89		
			(74.77)		
Pseudo-Thresh (85)				-51.69	
				(134.43)	
Pseudo-Thresh (90)					-145.64
					(193.79)
Observations	954	1,666	2,136	597	495
Mean Dep. Var.	30.60	68.47	96.37	235.87	280.99
Sample Empl. Range:	1 to 9	1 to 19	1 to 29	64 to 106	69 to 111
Panel B: Uniform Weights and Baseline Bandwidth					
Pseudo-Thresh (5)	-0.32				
	(49.74)				
Pseudo-Thresh (10)		-99.32**			
		(45.91)			
Pseudo-Thresh (15)			-43.64		
			(57.28)		
Pseudo-Thresh (85)				-47.81	
				(135.73)	
Pseudo-Thresh (90)					-130.78
					(192.30)
Observations	2,017	2,213	2,406	597	495
Mean Dep. Var.	79.81	80.54	84.18	227.02	280.85
Sample Empl. Range:	1 to 26	1 to 31	1 to 36	64 to 106	69 to 111

Notes: Dependent variable is R&D expenditures. Results from estimating the RDD model using placebo thresholds at which no discontinuities should exist, and indeed, no discontinuities are detected. In Panel A, triangular weights are used and an even bandwidth for each threshold going up to the baseline of 21 when possible, and in Panel B, uniform weights are used and the baseline bandwidth when possible. All baseline running variable controls, fixed effects, and additional controls included in all regressions. Standard errors are clustered by firm. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.8: No Difference-in-Discontinuities in Years Prior to Winning First Grant

Dep. Var.:	R&D (1)	R&D (2)	R&D (3)	R&D (4)
1[Empl. < 50] * Post 2012	-0.55	43.35	25.05	-156.56
	(79.76)	(91.76)	(97.56)	(175.57)
1[Empl. < 50]	86.51	24.48	42.26*	21.49
	(66.66)	(20.64)	(25.04)	(70.55)
Observations	1,281	1,279	1,136	951
Mean Dep. Var.	72.61	72.74	71.89	60.63
Baseline Controls		x	x	x
Firm FEs			x	x
Year x Industry FEs				x

Notes: Dependent variable is R&D expenditures (£000s). Table provides results from estimating the diff-in-disc model of Equation **??** for small firms when including only observations prior to when firms receive their first Innovate UK grant (i.e., when there should be no such discontinuities or differences in discontinuities). No statistically significant discontinuities or differences are detected. Baseline sample inclusion criteria are applied such that firms with 29 to 71 employees are included and only those that also meet the turnover and total assets grant generosity criteria. Running variable controls are included in all regressions and other baseline controls and fixed effects are included in Columns 2-4. Standard errors are clustered by firm. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.9: Event Study Point Estimates for Discontinuities in R&D by Year

Year (1)	Coefficient Estimate (2)	Standard Error (3)	P> |t| (4)
2005	298.33	474.89	0.530
2006	238.73	450.06	0.596
2007	282.29	450.64	0.532
2008	135.41	467.10	0.772
2009	339.62	425.27	0.425
2010	369.25	513.16	0.472
2011	231.25	300.32	0.442
2013	880.52	389.26	0.024
2014	573.44	522.47	0.273
2015	134.31	460.55	0.771
2016	442.17	458.25	0.335
2017	652.49	470.89	0.167

Average Discontinuities in Low vs. High Tax Credit Rate Periods

Low Tax Credit Rate Period (2005-12):	270.70
High Tax Credit Rate Period (2013-17):	536.59

Notes: Table provides point estimates, standard errors, and p-values for the coefficients from an event study version of Equation 1 estimating independent effect of more generous grant funding in R&D expenditures (£000s). The estimates are also plotted in Appendix Figure B.1. Coefficients capture the discontinuity in R&D at the grant generosity threshold each year conditional on the full baseline set of controls, running variables, and fixed effects. The model is estimated as one equation by interacting the grant generosity treatment indicator (and the running variables that differ at the threshold) with indicators for each year, omitting 2012 as the reference year. Standard errors are clustered by firm. Asterisks denote *$p<0.10$, **$p<0.05$, ***$p<0.01$.

Table C.10: Policy Interaction Effects for Small Firms When Widening Bandwidth

Dependent Variable:	R&D (1)	R&D (2)	R&D (3)	R&D (4)	R&D (5)	R&D (6)	R&D (7)
1[Empl. < 50] * Post 2012	654.95***	573.74***	530.95***	544.41***	569.80**	529.16**	470.42*
	(224.97)	(200.08)	(189.32)	(190.82)	(224.90)	(249.65)	(259.83)
1[Empl. < 50]	-111.40	-55.70	-80.70	-96.47	-336.04	-113.61	-141.29
	(217.01)	(203.77)	(210.26)	(233.04)	(293.05)	(326.05)	(307.51)
Observations	1,460	1,737	2,105	2,428	2,862	2,944	2,970
Mean Dep. Var.	184.54	178.34	175.00	184.87	175.36	191.22	190.51
Sample Empl Range:	20 to 71	15 to 71	10 to 71	10 to 90	10 to 150	10 to 200	10 to 250
Baseline Controls	x	x	x	x	x	x	x
Firm FEs	x	x	x	x	x	x	x
Year x Industry FEs	x	x	x	x	x	x	x

Notes: Table provides diff-in-disc effects of the policy interactions on small firms' R&D expenditures (£000s) when widening the window around the grant rate threshold. Sample includes only firms that also meet the total assets and turnover grant generosity criteria. Standard errors are clustered by firm. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.11: No Evidence of Expenditure Relabelling Expenditures

Dependent Variable:	Tangible Assets	Investment (incl. depr.)	Investment (no depr.)
	(1)	(2)	(3)
1[Empl. < 50] * Post 2012	173.14	-109.81	-175.12
	(445.80)	(215.89)	(228.83)
1[Empl. < 50]	-511.52	155.26	219.87
	(528.54)	(181.30)	(181.99)
Observations	982	958	982
Mean Dep. Var.	1182.26	226.54	19.00

Notes: Table provides estimates from estimating the difference-in-discontinuities in ordinary investment to test whether firms appear to relabel expenditures. Dependent variables are tangible assets (Column 1) and non-R&D ordinary investment (including depreciation in Column 2 and not including depreciation in Column 3). Baseline estimation sample is used and each regression includes all baseline running variables, controls, and fixed effects. Standard errors are clustered at the firm level. Asterisks denote *$p<0.10$, **$p<0.05$, ***$p<0.01$.

Table C.12: Pseudo-Threshold Tests for Diff-in-Disc Policy Interaction Effects

Dependent Variable:	R&D	R&D	R&D	R&D	R&D
	(1)	(2)	(3)	(4)	(5)
Panel A: Triangular Weights and Even Bandwidth					
Pseudo-Thresh (5) * Post 2012	-91.99				
	(84.11)				
Pseudo-Thresh (10) * Post 2012		-140.07			
		(179.70)			
Pseudo-Thresh (15) * Post 2012			171.78		
			(179.57)		
Pseudo-Thresh (85) * Post 2012				-232.70	
				(369.67)	
Pseudo-Thresh (90) * Post 2012					-174.66
					(595.92)
Observations	641	1,259	1,694	562	460
Mean Dep. Var.	38.23	84.13	112.57	233.22	291.48
Empl. Range:	1 to 9	1 to 19	1 to 29	64 to 106	69 to 111
Panel B: Uniform Weights and Baseline Bandwidth					
Pseudo-Thresh (5) * Post 2012	-129.84				
	(197.94)				
Pseudo-Thresh (10) * Post 2012		18.87			
		(180.66)			
Pseudo-Thresh (15) * Post 2012			48.59		
			(277.40)		
Pseudo-Thresh (85) * Post 2012				-272.94	
				(375.58)	
Pseudo-Thresh (90) * Post 2012					-175.93
					(622.53)
Observations	1,573	1,763	1,946	562	460
Mean Dep. Var.	97.69	96.62	98.74	223.61	291.78
Empl. Range:	1 to 26	1 to 31	1 to 36	64 to 106	69 to 111

Notes: Dependent variable is R&D expenditures. Results from estimating the diff-in-disc model using placebo thresholds at which no discontinuities should exist. No differences in discontinuities are detected, as expected. In Panel A, triangular weights are used and an even bandwidth for each threshold going up to the baseline of 21 when possible, and in Panel B, uniform weights are used and the baseline bandwidth when possible. All baseline running variable controls, fixed effects, and additional controls included in all regressions. Standard errors are clustered by firm. Asterisks denote $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Table C.13: Patent Box and R&D Tax Credit Claims

	R&D Tax Credit Relief		**Patent Box Relief**			
Fiscal Year	SME Total Claims (1)	SME Total Relief (£m) (2)	Small Claims (3)	Medium Claims (4)	SME Total Claims (5)	SME Total Relief (£m) (6)
2004/05	5,310	190				
2005/06	4,960	185				
2006/07	5,270	200				
2007/08	5,990	245				
2008/09	6,670	265				
2009/10	7,470	320				
2010/11	8,280	355				
2011/12	10,030	435				
2012/13	13,140	615				
2013/14	15,585	705	170	175	345	14.8
2014/15	29,775	1,315	285	275	560	32.9
2015/16	37,105	1,760	280	285	565	31.2
2016/17	45,440	2,265	265	260	525	32.7
Averages (2013-17)	**31,976**	**1,511**	**250**	**249**	**499**	**27.9**

Notes: Table provides summary statistics of R&D Tax Credit Scheme claims for SMEs (Columns 1-2) and Patent Box claims (Columns 3-6). Data are compiled by author from annual Patent Box and R&D Tax Credit official statistics published on the HM Revenue and Customers (HMRC) website.

Table C.14: Robustness to Omitting Firms in Sectors with Many Patent Box Claims

Dependent Variable:	R&D (1)	R&D (2)	R&D (3)
1[Empl. < 50] * Post 2012	466.88*	439.65**	708.00**
	(242.49)	(187.13)	(332.79)
1[Empl. < 50]	-97.30	-81.84	-675.58*
	(68.15)	(83.24)	(371.17)
Observations	740	721	536
Mean Dep. Var.	197.29	199.41	237.08
Baseline Controls		x	x
Firm FEs			x
Year x Industry FEs			x

Notes: Dependent variable is R&D expenditures (£000s). Firms in sectors with majority of patent box claims (manufacturing and wholesale and trade sectors) are omitted and sample is otherwise the same as the baseline, including firms with 29 to 71 employees in the year before winning a grant conditional on also meeting the turnover and total assets grant rate generosity threshold. Running variable controls included in all regressions. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.15: Robustness of Policy Interaction Effects to Different Winsorization Rules and Addressing R&D Reporting Concerns

Dependent Variable:	R&D (1)	R&D/Empl. (2)	R&D/SIC Total (3)	R&D (4)	R&D (5)
1[Empl. < 50] * Post 2012	507.69**	10.57*	0.11**	627.78**	687.27**
	(244.30)	(5.82)	(0.05)	(250.51)	(302.20)
1[Empl. < 50]	-192.84	-2.71	-0.05	-95.02	-328.79
	(226.04)	(6.12)	(0.07)	(282.20)	(321.45)
Observations	1,043	974	1,047	877	717
Mean Dep. Var.	157.57	3.53	0.04	218.60	179.95
Firm FEs	x	x	x	x	x
Year x Industry FEs	x	x	x	x	x
Additional Controls	x	x	x	x	x
Winsorization	5%	1%	1%	1%	1%
Non-Missing Profit Data				x	
Non-Missing Cost of Sales Data					x

Notes: Dependent variable is R&D expenditures (£000s) in Columns 1, 4, and 5. In Column 2, the dependent variable is R&D per employee, and in Column 3, it is R&D as a proportion of the firm's 4-digit SIC total. In Column 1, I winsorize the top 5% of the distribution of non-zero R&D expenditures (for firms with fewer than 100 employees) rather than 1% as done in the baseline. In Columns 4 and 5, I limit the sample to only firms with non-missing data for profits and cost of sales variables, respectively. Sample includes firms with 29 to 71 employees in the year before winning a grant conditional on also meeting the turnover and total assets grant rate generosity eligibility criteria. All baseline running variables, fixed effects, and additional controls included in all regressions. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.16: Robustness of Policy Interaction Effects to Different Decisions Determining Sample Selection

Dependent Variable:	R&D (1)	R&D (2)	R&D (3)	R&D (4)	R&D (5)	R&D (6)	R&D (7)
1[Empl. < 50] * Post 2012	540.60**	506.28**	562.07**	519.14**	769.14***	781.37**	924.01**
	(260.48)	(227.48)	(222.71)	(236.26)	(268.79)	(339.78)	(385.83)
1[Empl. < 50]	-391.66	76.28	-121.37	-123.29	-390.64	-393.70	-840.24*
	(353.58)	(226.24)	(259.61)	(293.48)	(361.80)	(426.61)	(502.84)
Observations	927	1,324	985	1,181	930	863	803
Mean Dep. Var.	199.41	211.27	195.71	181.50	210.54	220.38	226.69
Years in Sample:	2005-17	2005-17	2005-17	2005-17	2008-17	2009-17	2010-17
Firm FEs	x	x	x	x	x	x	x
Year x Industry FEs	x	x	x	x	x	x	x
Additional Controls	x	x	x	x	x	x	x

Notes: Dependent variable is R&D expenditures (£000s). In Column 1, the bandwidth is reduced to 19 such that the sample includes only firms with 31 to 69 employees in the year prior to receiving a grant (but still conditional on meeting the other eligibility criteria). In Column 2, I use the baseline MSE-optimal bandwidth but do not condition on firms also meeting the other grant rate generosity criteria. In Column 3, I omit observations associated with grants received after 2015 so there are at least 2 years of post-grant data. In Column 4, I include up to four years of data post-grant rather than just three. In Columns 5-7, I omit years leading up to the Great Recession. The estimation sample is otherwise the same as the baseline, including firms with 29 to 71 employees in the year before winning a grant conditional on also meeting the turnover and total assets grant rate generosity eligibility criteria. Running variable controls included in all regressions. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.17: Additional Robustness Checks for Policy Interaction Effects on Small Firms' R&D Expenditures

Dependent Variable:	R&D Cluster by Industry (1)	R&D Uniform Weights (2)	R&D Quadratic Polynomials (3)	R&D Cubic Polynomials (4)
1[Empl. < 50] * Post 2012	560.68**	490.77**	711.84***	746.52***
	(229.61)	(237.37)	(233.49)	(273.26)
1[Empl. < 50]	-122.12	94.82	-178.03	214.47
	(312.86)	(232.06)	(385.45)	(639.24)
Observations	1,046	1,324	1,046	1,046
Mean Dep. Var.	186.89	212.16	186.89	186.89
Polynomial Flexibility:				
Linear (baseline)	x	x		
Quadratic			x	
Cubic				x
Firm FEs	x	x	x	x
Year x Industry FEs	x	x	x	x
Additional Controls	x	x	x	
Clustering:	Industry	Firm	Firm	Firm
Kernel Weight:	Tri.	Uni.	Tri.	Tri.

Notes: Dependent variable is R&D expenditures (£000*s*). Sample includes firms with 29 to 71 employees in the year prior to winning a grant conditional on also meeting the turnover or total assets eligibility criteria for more generous grant funding rates. Column 1 clusters standard errors at the 4-digit SIC level rather than firm level. In Column 2, I remove the requirement for the sample to only include firms that also meet the turnover and total assets grant rate generosity threshold. This sample selection criteria apply again though in Columns 3-5. In Column 3, I use uniform weights rather than triangular, and in Columns 4 and 5, I increase the polynomial flexibility of the running variable controls. Standard errors are clustered at the firm level in Columns 2-5. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.18: Details on Financial Constraint Proxy Variables

Proxy Variable	Description	Median in Baseline Sample (£000s)
Operating Profit	Gross profit minus operating expenses before accounting for net interest paid/received and taxes Constrained firms are more likely to have lower operating profits. Variable is constructed by author using FAME data.	£69
Available Funds	After-tax profits plus depreciation Constrained firms are more likely to have fewer available funds. Variable is taken directly from FAME database.	£174
Short-term Loans	Current liabilities such as group loans, bank overdrafts, short-term hire purchasing and leasing, etc. Constrained firms are more likely to have more short-term debt. Variable is taken directly from FAME database.	-£411

Notes: Table provides definitions of financial constraint proxies used when estimating the heterogeneous policy effects for constrained vs. unconstrained firms in Section 7.1.1. The proxies aim to capture the resources firms may have for self-financing R&D. Median values in the year prior to when firms in the baseline estimation sample receive a grant are also included, which I use when splitting the sample into constrained versus unconstrained firms. Results from the heterogeneous analyses are in Table 5.

Table C.19: Pre-Policy Covariate Balance for Larger Firms

Dep. Var.:	Grant Funds (£000s) (1)	Grant Funds per R&D Exp. (2)	Revenue (£m) (3)	Revenue per Empl. (4)	Age (years) (5)	Average Wages (6)	Number of Scientists (7)
1[Empl.<500]	2.67	0.01	-9.79	-31.57	-0.52	0.16	-0.49
	(5.64)	(0.01)	(9.83)	(28.26)	(0.86)	(0.33)	(2.29)
Observations	2,746	2,746	2,746	2,746	2,746	2,746	2,746
No. of Firms	1299	1299	1299	1299	1299	1299	1299

Notes: Table provides evidence of balanced covariates around the tax credit policy threshold in the pre-policy period (2000-2007), suggesting that firms around the threshold are similar. The main regressor is a dummy variable equal to one if the firm has fewer than 500 employees within a regression discontinuity design. Firms with 250 to 750 employees are included in the sample. In all cases, triangular weights are used and first-order polynomials of the running variable (employment) are included separately for each side of the threshold. In Column 1, the dependent variable is direct subsidies for R&D (000s GBP). In Column 2, it is the proportion of R&D expenditures subsidized by government subsidies. In Column 3, it is revenue (millions GBP). In Column 4, it is labor productivity (turnover in millions over total number of employees). In Column 5, it is firm age. In Column 6, it is average wages (R&D worker salaries in total over total number of employees). In Column 7, it is the number of R&D scientists. The top 1% of pre-policy R&D expenditure distribution is dropped to account for outliers in all cases. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.20: Robustness of Tax Credit Generosity Effects to Different Windows Around Threshold for Larger Firms

Sample Empl. Range:	300-700	275-725	225-775	200-800
Dependent Variable:	R&D	R&D	R&D	R&D
	(1)	(2)	(3)	(4)
Panel A: Discontinuities in R&D in Post-Policy Period (2009-2014)				
1[Empl.<500]	759.80*	924.11**	1,021.35**	999.13**
	(452.19)	(441.02)	(441.99)	(434.71)
Observations	1,100	1,253	1,528	1,702
No. of Firms	466	538	635	704
Dep. Var. Mean	2920	2862	2828	2799
Panel B: Discontinuities in R&D in Pre-Policy Period (2000-2008)				
1[Empl.<500]	230.80	221.26	136.15	133.46
	(163.25)	(159.94)	(149.83)	(145.15)
Observations	2,141	2,446	3,063	3,425
No. of Firms	1040	1155	1428	1603
Dep. Var. Mean	1809	1791	1754	1738

Notes: Table provides results from estimating the effects of being eligible for more generous tax credits on larger firms' R&D expenditures (£000s) using the RDD model but alternative firm size bandwidths around the 500-employee threshold. The windows of firm sizes are provided in the header. Dependent variable is R&D expenditures (£000s). Panel A provides the discontinuities in the post-policy period (2009-2014) and Panel B provides the discontinuities in the pre-policy period (2000-2008). In all regressions, triangular weights are used and first-order polynomials of the running variable (employment) are included separately for each side of the threshold. All regressions in Panel A also include a control for the firm's average pre-policy R&D expenditures and I winsorize by dropping the top 1% of the R&D expenditure distribution. Standard errors are clustered at the firm level. Asterisks denote *$p<0.10$, **$p<0.05$, ***$p<0.01$.

Table C.21: Falsification Tests for Policy Interaction Effects for Larger Firms

Dependent Variable:	R&D	R&D	R&D	R&D	Non-Cap R&D	Cap R&D
Threshold Type::	Pseudo	Pseudo	Pseudo	Pseudo	Real	Real
	(1)	(2)	(3)	(4)	(5)	(6)
Direct Subsidies * Threshold	0.46	0.33	1.94	4.02	-2.64**	-0.19
	(0.6)	(1.95)	(2.71)	(4.21)	(1.34)	(0.12)
Direct Subsidies	2.42***	2.63***	6.19***	6.46**	3.62***	0.22**
	(0.23)	(0.34)	(1.56)	(2.96)	(1.18)	(0.11)
Observations	7,869	3,405	1,646	1,035	1,382	1,382
No. of Firms	5086	1768	719	450	584	584
Dep. Var. Mean	647.6	1998	2531	2897	2632	179.5
Threshold (Empl.):	100	350	650	900	500	500
Sample Empl. Range:	0-200	100-600	400-900	650-1150	250-750	250-750

Notes: Table provides results from falsification tests of larger firm policy interaction effects when estimating discontinuity-in-effects models. In Columns 1-4, the dependent variable is R&D expenditures (*£000s*) and I conduct placebo tests imposing fake pseudo-thresholds, finding no statistically significant discontinuities. In Columns 5 and 6, I return to using the actual tax credit generosity threshold and estimate the effects specifically on non-capital R&D (Column 5), where the substitution is most likely to occur, and capital R&D (Column 6), where there should be less or no substitution since these expenditures typically do not qualify for tax credits. In all regressions, triangular weights are used and first-order polynomials of the running variable (employment) are included separately for each side of the threshold. All regressions also include a control for the firm's average pre-policy R&D expenditures and I winsorize by dropping the top 1% of the R&D expenditure distribution in the post-policy period. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.22: Using Lagged Employment to Define Tax Policy Treatment Status

Dep. Var.:	R&D	R&D
Employment year(s) used to define tax credit treatment:	Current + One Year Lag	Current + Two Year Lags
	(1)	(2)
Direct Subsidies * 1[Empl.<500]	-2.28*	-2.25*
	(1.30)	(1.29)
Direct Subsidies	3.38***	3.38***
	(1.11)	(1.11)
Observations	1,382	1,380
No. of Firms	584	583

Notes: Dependent variable is R&D expenditures (£000*s*). Table provides results from placing more stringent requirements on how tax credit eligibility is defined for firms around the 500-employee threshold by using one year of lagged employment in addition to the current year in Column 1 and two years of lagged employment plus the current year in Column 2. In both regressions, firms with 250 to 750 employees are included, triangular weights are used, and first-order polynomials of the running variable (employment) separately for each side of the threshold are included as well as a control for the firm's average pre-policy R&D expenditures. I drop the top 1% of the R&D expenditure distribution to account for outliers. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.

Table C.23: Additional Robustness Checks of Policy Interaction Effects on Larger Firms' R&D Exepnditures

Dependent Variable:	R&D (1)	R&D (2)	R&D (3)	R&D (4)
Panel A: Disc-in-Effects in Post-Policy Period (2009-2014)				
Direct Subsidies * 1[Empl.<500]	-2.86**	-2.87**	-3.36**	-3.64***
	(1.41)	(1.42)	(1.41)	(1.40)
Direct Subsidies	3.85***	3.86***	4.49***	5.04***
	(1.24)	(1.25)	(1.26)	(1.22)
1[Empl.<500]	653.91	562.40	1,244.51***	643.17***
	(509.54)	(544.69)	(465.30)	(211.11)
Observations	1,382	1,382	1,382	1,078
Dep. Var. Mean	2,811	2,811	2,766	1,453
Panel B: Disc-in-Effects in Pre-Policy Period (2000-2008)				
Direct Subsidies * 1[Empl.<500]	-0.42	-0.38	-0.22	1.69
	(2.31)	(2.31)	(2.45)	(2.38)
Direct Subsidies	3.23**	3.27**	3.13***	5.83***
	(1.29)	(1.30)	(1.20)	(1.56)
1[Empl.<500]	155.95	814.57	335.31	-12.29
	(475.24)	(626.37)	(315.36)	(143.45)
Observations	2,746	2,746	2,746	2,269
Dep. Var. Mean	1766	1766	1751	970.9
Polynomial Flexibility:				
Linear			x	x
Quadratic	x			
Cubic		x		
Only if Subsidies>0				x
Kernel Weighting:	Triangular	Triangular	Uniform	Triangular

Notes: Table provides robustness checks of the policy interaction effects for larger firms when implementing the discontinuity-in-effects approach. Dependent variable is R&D expenditures (£000s). Panel A provides estimates for the post-policy period (2009-2014) when discontinuities are expected and Panel B provides estimates for the pre-policy period (2000-2008) where a positive correlation between direct subsidies and R&D is expected but no discontinuity. In Columns 1 and 2, I increase the flexibility of the running variable controls. In Column 3, I estimate the baseline model using first-order polynomials but apply uniform kernel weighting rather than triangular. In Column 4, I limit the sample to include only observations for which there is a positive value of direct subsidies. In all regressions, firms with 250 to 750 employees are included and the top 1% of the R&D expenditure distribution is omitted to account for outliers. Standard errors are clustered at the firm level. Asterisks denote *$p<0.10$, **$p<0.05$, ***$p<0.01$.

Table C.24: No Evidence that Effects are Driven by Changes in Type of Projects Funded, Competition, or Cumulative Funding

Dependent Variable:	Feasibility (1)	Concept (2)	Prototype (3)	Market (4)	R&D (5)	R&D (6)
1[Empl. < 50] * Post 2012	-0.08	-0.02	-0.13	-0.04	1009.35***	964.38***
	(0.14)	(0.06)	(0.09)	(0.05)	(384.39)	(280.29)
1[Empl. < 50]	-0.07	0.01	0.08	0.06	2815.61***	-138.45
	(0.11)	(0.03)	(0.07)	(0.04)	(729.34)	(282.21)
1[Empl. <50] * Post 2012 * Cumulative Grants						-129.55*
						(67.73)
1[Empl. <50] * Cumulative Grants						61.93
						(71.04)
Cumulative Grants						-53.37
						(52.96)
Observations	1,047	1,047	1,047	1,047	971	1,047
Mean Dep. Var.	0.13	0.02	0.08	0.03	162.74	186.70
Firm FEs	x	x	x	x	x	x
Year x Industry FEs	x	x	x	x	x	x
Baseline Controls	x	x	x	x	x	x
Grant Competition FEs					x	

Notes: Table provides results from investigating alternative explanations of positive interaction effects and are associated with discussion in Section 7.2. In Columns 1-4, the dependent variables are indicators for whether projects are feasibility studies (Column 1), proofs of concept (Column 2), development of prototypes (Column 3), and proofs of concept (Column 4). In Column 5, I estimate the baseline model with grant competition-level fixed effects, and in Column 6, the main treatment variables are interacted with the firm's cumulative number of grants. Sample includes firms with 29 to 71 employees in the year prior to winning a grant conditional on also meeting the turnover and total assets grant rate generosity criteria. All baseline running variables, controls, and fixed effects are included in all regressions. Standard errors are clustered at the firm level. Asterisks denote *p <0.10, **p <0.05, ***p <0.01.