Economics of Innovation and New Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gein20

Information asymmetry, risk aversion and R&D subsidies: effect-size heterogeneity and policy conundrums

Mehmet Ugur & Eshref Trushin

To cite this article: Mehmet Ugur & Eshref Trushin (2022): Information asymmetry, risk aversion and R&D subsidies: effect-size heterogeneity and policy conundrums, Economics of Innovation and New Technology, DOI: 10.1080/10438599.2022.2119563

To link to this article: https://doi.org/10.1080/10438599.2022.2119563

Published online: 07 Sep 2022.

ECONOMICS OF INNOVATION AND NEW TECHNOLOGY
https://doi.org/10.1080/10438599.2022.2119563

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Information asymmetry, risk aversion and R&D subsidies: effect-size heterogeneity and policy conundrums

Mehmet Ugur [a] and Eshref Trushin [b]

[a]Institute of Political Economy, Governance, Finance and Accountability, University of Greenwich, London, UK; [b]Department of Economics, De Montfort University, Leicester, UK

ABSTRACT
Drawing on the theory of contracts and Schumpeterian models of innovation, we demonstrate that information asymmetry and risk aversion are conducive to effect-size heterogeneity and sub-optimal allocation of R&D subsidies. Utilising an unbalanced panel of 43,650 British firms from 1998 to 2012 and an entropy balancing methodology, we find that R&D subsidies are less likely to generate additionality effects when: (a) firms are larger, older, or more R&D-intensive; and (b) investment in basic research or during crisis episodes is considered. We also report that over 85% of the subsidies are allocated to large, old and R&D-intensive firms that do not deliver additional R&D investment. Our findings reveal a policy conundrum: the case for R&D subsidies is stronger during economic downturns, when R&D investment is in basic research and when firm age, size and R&D intensity reflect success in converting R&D investment into innovative product lines; but the subsidy is less likely to increase business R&D under these conditions.

ARTICLE HISTORY
Received 25 May 2022
Accepted 27 August 2022

KEYWORDS
R&D subsidy; additionality; information asymmetry; risk version; entropy balancing

JEL Classification
O31; O38; C21

1. Introduction

Since the early review by David, Hall, and Toole (2000), the number of studies on the effectiveness of research and development (R&D) subsidies has increased exponentially, reflecting both the wide-spread adoption of the support schemes internationally and the developments in estimation methods (for reviews, see Becker 2015; Cerulli 2010; Dimos and Pugh 2016; Zúñiga-Vicente et al. 2014). As observed by Cerulli (2010, 424), most of the studies have tended to focus on whether public support has been associated with input additionality '… skipping, at least implicitly, the essential step of going into an explicit theoretical framework to explain this causal relation[ship]'.

This paper is concerned with another 'under-theorised' aspect of the work on the effectiveness R&D subsidies: the extent of effect-size heterogeneity and its potential sources that relate to firm characteristics and R&D types. Indeed, the existing evidence syntheses indicate that R&D subsidies do not generate input additionality or are associated with crowding-out effects in about 40–45 percent of the effect-size estimates reported in primary studies. Moreover, the estimates in the remaining 55–60 percent of the evidence base differ by firm age, size or R&D intensity and by R&D type such as basic, applied or capital R&D. This is despite expectations of additionality effects informed by the 'market failure' and/or 'financing constraint' arguments that inform public policy.

CONTACT Mehmet Ugur m.ugur@gre.ac.uk Institute of Political Economy, Governance, Finance and Accountability, School of Accounting, Economics and Finance, University of Greenwich, Park Row, London SE10 9LS, UK
Supplemental data for this article can be accessed online at https://doi.org/10.1080/10438599.2022.2119563.

This state of affair has led Zúñiga-Vicente et al. (2014) to take stock of the existing evidence and put forward a range of propositions about the sources of observed heterogeneity. For example, the authors consider the firm's subsidy history (particularly, the frequency with which it receives subsidies), the extent to which it may be financially constrained and the type of R&D (e.g. applied *versus* basic R&D) investment it undertakes as the most relevant. These are followed by another set of factors that consists of the time lag between the subsidy and the effect-size measurement, the generosity of the subsidy rate and the multiplicity of the funding agencies (sources).

The work by Zúñiga-Vicente et al. (2014) is a welcome step in the right direction. The strength of the work derives from the authors' bottom-up approach – i.e. from their careful investigation and synthesis of the existing empirical patterns. Against this strength, however, the causal/theoretical explanations on offer remain inevitably eclectic, reflecting the *ex-post* and mostly partial/limited theoretical explanations already offered in the primary studies.

One aim of this paper is to provide a *unified theoretical framework* that extends the emerging explanations of effect-size heterogeneity by taking account of the risk aversion and information asymmetry that characterise the principal-agent relationship between the public funder and the eligible firms. To do this, we draw on the theory of contracts (Laffont and Martimort 2002; Akcigit, Hanley, and Stantcheva 2019), where the two causal sources of heterogenous subsidy effects are *risk aversion* of the eligible firms and *information asymmetry* between the latter and the public funder. Whilst risk aversion attenuates the effect of the R&D subsidy as the perceived level of return uncertainty increases, information symmetry about the eligible firms' R&D productivity enables high-R&D-productivity firms to extract informational rents and exert less effort per unit of subsidy received.

Another aim of our paper is to map the *unobserved* risk aversion and information asymmetry on to observable firm and R&D characteristics that would be associated with variation in additionality effects. To achieve this aim, we combine the insights from the theory of contracts with key predictions from Schumpeterian models of innovation (Aghion, Akcigit, and Howitt 2014, 2015; Akcigit, Hanley, and Stantcheva 2019; Strulik 2007). The combination enables us to derive theory-informed hypotheses on the causal links between observed effect-size heterogeneity and its observed sources (i.e. observed firm and R&D characteristics) reported either in individual empirical studies or in reviews thereof. The proposed hypotheses can be summarised as follows: (i) the effects of R&D subsidies on basic research or during crisis periods would be smaller or insignificant; and (ii) R&D subsidies are less likely to generate input additionality effects when the subsidised firms are older, larger and have higher levels of R&D intensity. In (i), the prediction of smaller or insignificant effects is due to higher levels of perceived return uncertainty in basic research or during crisis episodes. In (ii), it is due to narrower R&D gaps (i.e. smaller difference between actual and privately optimal levels of R&D investment) among old, large and R&D-intensive firms – which would have higher R&D productivity and would invest in R&D with or without the subsidy.

The third aim of our paper is to adopt a relatively novel empirical approach based on entropy balancing (EB) (Hainmueller 2012).[1] A major advantage of the EB estimator is that it does not depend on a propensity score model, the correct specification of which is a major challenge. Also, it ensures a high degree of covariate balancing between treatment and control groups along three moment conditions: mean, variance, and skewness. We apply the EB methodology to an unbalanced panel of 43,650 British firms from 1998 to 2012 and provide a range of robustness checks to verify the stability of our findings.

In the empirical part of the paper, we report that R&D subsidies generate additionality effects among small and young firms and firms away from the R&D frontier that account only for about 2–15 percent of the subsidy allocations. In contrast, R&D subsidies have no significant effects on R&D investment of large and old firms and firms closer to the R&D frontier in the industry, which account for more than 80 percent of the subsidy allocations. Furthermore, the treatment effect is larger when the investment is in applied or experimental R&D, where the non-appropriability problem is less severe due to higher market-readiness of the R&D type. In contrast, the additionality effect is smaller or insignificant during crisis episodes or when investment is in basic research, where

R&D subsidies are expected to ameliorate the financing constraints and/or correct for the dis-incentive effects of non-appropriability and return uncertainty.

Our findings, which remain robust to a range of robustness checks, point out to a policy conundrum in public support for business R&D. On the one hand, it is socially optimal to provide public support for basic research, during economic downturns, and when firms have a history of success in converting R&D inputs into profitable innovations. On the other hand, R&D subsidies are less likely to generate additionality effects under these conditions.

The rest of the paper is organised as follows. In section 2, we review the empirical literature and demonstrate that the reported effect-size estimates differ between firm types, R&D investment types, and phases of the business cycle. We also demonstrate that the explanations for effect-size heterogeneity have remained largely fragmented and unsystematic. Drawing on these observations, we indicate the need for a unifying theoretical framework that would allow for testable hypotheses/predictions about the sources of observed heterogeneity.

The unifying theoretical framework is developed in sections 3 and 4. In section 3, we first highlight the main feature of the R&D subsidy regime that governs the principal-agent relationship between the public funders in Great Britain and the eligible firms that constitute our sample. This is followed by a discussion on the implications of risk aversion and information asymmetry in the theory of contracts and mechanism design literature (Akcigit, Hanley, and Stantcheva 2019; Laffont and Martimort 2002; Strulik 2007). The analysis in section 3 enables us to demonstrate that subsidy allocations would remain suboptimal and firm responses to R&D subsidies would be inherently heterogenous if: (i) firms are risk averse (i.e. if they utilise higher discount rates for evaluating R&D projects when perceived return uncertainty increases); and (ii) information asymmetry exists between funded firms and the public funder with respect to the firms' R&D productivity, risk aversion, and R&D gaps.

In section 4, we map the *unobservable* information asymmetry, risk aversion and R&D gaps onto *observable* firm characteristics (e.g. firm age, size, and R&D intensity) and R&D types (e.g. basic *versus* applied R&D) to derive testable hypotheses about the observed sources of heterogeneity in the effects of R&D subsidies on business R&D. Hence, sections 3 and 4 constitute our unifying theoretical framework, which identifies the conditions under which the allocation of R&D subsidies might be sub-optimal and the firm/R&D types that might be associated with smaller or insignificant additionality effects.

In section 5, we present our dataset, introduce the entropy balancing (EB) methodology we use to identify causal effects, and discuss the range of robustness checks we conduct to address issues such as multiple treatments and omitted unobservables. The estimation results are presented in section 6, complemented with additional robustness and balancing checks in the on-line Appendix. Finally, section 7 concludes.

2. Effect-size heterogeneity in the empirical literature: what do we know?

The extent of heterogeneity in the effects of R&D subsidies on business R&D is well documented (David, Hall, and Toole 2000; García-Quevedo 2004; Cerulli 2010; Zúñiga-Vicente et al. 2014; Becker 2015; Dimos and Pugh 2016; Antonelli 2020).[2] This is not surprising given the contingent and somewhat open-ended nature of the theoretical case for public support of business R&D, which depends on market failures that generate an *R&D gap* between a firm's actual R&D investment and the socially optimal level of R&D. One source of market failure is incomplete appropriability of innovation outcomes/benefits (Arrow 1996; Romer 1990). Another relates to capital-market imperfections, which exacerbate the financing constraints of the innovative firms with inadequate collateral (Bloom, Bond, and van Reenen 2007; Czarnitzki and Toole 2007; Hall 1992, 2002, 2010; Minton and Schrand 1999).

Nevertheless, market failures may generate different R&D gaps for several reasons. First, the investment-deterring effects of market failures differ between firms, depending on the variation in the latter's size, market power and ability to protect their intellectual property assets (Antonelli 2020). Secondly, the investment-deterring effects of knowledge spillovers unfold in the context of intra-industry competition, which induces firms to invest in own R&D to minimise the market-stealing effects of R&D investment by competitors (Bloom, Van Reenen, and Williams 2019). The R&D gap may also differ between firms because of the need to invest in R&D before a firm can benefit from knowledge spillovers (Antonelli 2020; Branstetter 1998; Cohen and Levinthal 1989; Geroski 1995).[3]

Hence, it is safe to conclude that the firms' R&D gaps, hence their responses to R&D subsidies, are inherently heterogenous – depending on the firm-specific balance between the *investment-deterring effects* of market failures and the *investment-inducing effects* of R&D productivity and intra-industry competition. Given this scope for variation in R&D gaps, the earliest pattern of effect-size heterogeneity has been reported by David, Hall, and Toole (2000). The authors review 14 studies, of which three reported additionality effects, five studies reported crowding-out effects, and the remaining six reported mixed findings. Here, effect-size heterogeneity reflects a geographical dimension: whereas studies based on US data are more likely to report insignificant or crowding-out effects, those based on European data tend to report additionality effects.

Another early review is by García-Quevedo (2004), who conducts a meta-analysis of 74 studies at the firm, industry, and country levels. The vote-counting exercise in this review indicates that 38 studies report additionality effects whereas 36 studies report insignificant (19) or crowding-out (17) effects. To identify the sources of heterogeneity, García-Quevedo (2004) controls for a range of moderating variables, including the level of analysis (firm, industry, or country), data structure (panel versus cross-section data), whether the model allows for lagged effects, and United States (US) *versus* non-US data, etc. Although the descriptive evidence indicates higher probability of crowding-out effect when US data is used and the analysis is at the firm level, this is not supported by the formal analysis of a general meta-regression model.

The third review by Cerulli (2010) examines model specifications and estimation methods as potential sources of effect-size heterogeneity. The review reports that the estimation methods that control for subsidy endogeneity (e.g. selection or instrumental variable treatment-effect models) and those that control for both observables and unobservable factors (e.g. difference-indifference) yield *smaller* additionality effects compared to structural models without selection or those that control for observables only (e.g. matching or propensity-score weighting). A second conclusion in the review relates to the need for explicit theorisation of the market structure and the macro-economic environment as significant contextual factors that affects the firms' expectation formation, their R&D investment decisions in general, and their responses to R&D subsidies in particular.

A meta-analysis of 660 effect-size estimates from 52 firm-level studies (Dimos and Pugh 2016) control for both estimation methods and firm/R&D project types as potential sources of heterogeneity – in addition to data type and year of publication. The authors report that studies based on small-firm data and high-tech firm samples tend to report *smaller* input additionality. Input additionality is also *smaller* and may indicate *crowding-out* effects when the estimation method controls for endogeneity and for both observable and unobservable factors (e.g. difference-in-difference). Overall, the study-specific average effects are insignificant or indicate crowding out in 45% of the evidence base.

A systematic review of 168 effect-size estimates from 77 studies by Zúñiga-Vicente et al. (2014) confirms some of the findings discussed so far. For example, 60% of the reviewed findings indicate input additionality, while the remaining 40% of the evidence base indicates crowding-out or insignificant effects. Moreover, the magnitude of the statistically significant input additionality effects varies between studies, by firm characteristics, R&D types, and subsidy regimes.

The added value of the review by Zúñiga-Vicente et al. (2014) is that it explicitly aims to develop testable postulates about the sources of effect-size heterogeneity, which include firm and R&D investment types, financial constraints, and the subsidy and funding history f the firm. The authors aim to establish which of these factors are more or less likely to be associated with input additionality effects. The authors report that the history of past subsidies and multiple founding sources have ambiguous effect on input additionality; whereas longer time lags and more severe financing constraints are more likely to be associated with input additionality effects. Hence, R&D subsidies to small and young firms and firms investing in basic research are more likely to be associated with input additionality.

Although the proposed postulates are helpful in classifying potential sources of heterogeneity into manageable categories, further work remains necessary. As already acknowledged by Zúñiga-Vicente et al. (2014), there are further research issues to be considered – for example the impact of economic crises. Secondly, and more importantly, the proposed postulates do not offer a unifying theoretical framework, which is needed to identify the causal channels through which firm and R&D characteristics would be conducive to higher or lower probabilities of input additionality.

The need for a unifying theoretical framework becomes more obvious when the role of financial constraints is considered. Because financing constraints result from failures in both innovation and capital markets, they are the most reported determinant of input additionality – in both individual studies and in reviews such as Zúñiga-Vicente et al. (2014). Indeed, input additionality among small and young firms have been related to financing constraint in several studies (Aristei, Sterlacchini, and Venturini 2017; González and Pazó 2008; Hassine and Mathieu 2020; Lach 2002; Nilsen, Raknerud, and Iancu 2018). In the same vein, insignificant or crowding-out effects have been explained by weaker financial constraints among larger firms (Czarnitzki and Hottenrott 2011; Szücs 2020; Boeing, Eberle, and Howell 2022).

Nevertheless, the financing constraints argument is necessary but not sufficient for explaining why R&D subsidies may or may not generate additionality effects. For example, the financing constraint is more severe in the case of basic R&D projects or when the investment is during crisis periods. However, there are conflicting findings about the effectiveness of R&D subsidies under these conditions. For example, Clausen (2009) and Czarnitzki, Hottenrott, and Thorwarth (2011) report that input additionality is more likely when the investment is in basic research. In contrast, Aerts and Thorwarth (2008) report crowding-out effects as the negative effect of return uncertainty in the case of basic research investment dominates the positive effect of the subsidy that ameliorates the financial constraint. Similar evidence has been reported with respect to the effects of R&D subsidies during crisis periods (Hud and Hussinger 2015).

Moreover, Antonelli (2020) questions the Arrovian argument for R&D subsidies by highlighting the limited exhaustibility of knowledge and the diachronic knowledge externalities. One argument in this paper is that the difference between private and social returns (hence, the difference between privately and socially optimal levels of R&D investment) would fall as firms become better able to appropriate the stream of benefits associated with innovation. The implication is that firms are likely to have different R&D gaps and would respond to R&D subsidies differently. A second argument is that R&D subsidies are more likely to be associated with input additionality when markets for both knowledge and goods production are competitive. If markets are characterised by imperfect competition, the R&D subsidies may fail to generate input additionality as the R&D investment of the firms with market power would be closer to the privately optimal level of R&D investment even without the subsidy. Hence, Antonelli (2020) concludes that innovation policy should move away from R&D subsidies aimed at reducing the cost of R&D to conditional R&D support schemes with mandatory provisions for increasing the flow of knowledge.

Given the 'state of the art' summarised above, the need for developing a unifying theoretical framework remains evident. We address this knowledge gap by combining insights from three sources: (i) the theory of contracts and mechanism design literature that analyse the R&D subsidy as a principal-gent problem under information asymmetry and risk aversion (Laffont and Martimort 2002; Salanie 2005; Akcigit, Hanley, and Stantcheva 2019).; (ii) the Schumpeterian models of innovation that analyse the firm's R&D intensity as an endogenous choice that, in turn, determines its size, age, and proximity to the R&D frontier (Aghion, Akcigit, and Howitt 2014, 2015); and (iii) the 'waiting' theory of investment that analyses firm expectations, their investment decisions, and their responsiveness to policy interventions under return uncertainty (Bernanke 1983; Bloom, Bond, and van Reenen 2007; Bouvatier and Lepetit 2008).

The overlapping insights from these theoretical perspectives will enable us to develop a unifying theoretical framework that explains the potential for sub-optimal subsidy allocations and effect-size heterogeneity by: (i) taking account of market failures, financing constraints and the principal-agent relationships in the subsidy regime at the same time; (ii) identifying the *observable* firm types in terms of age, size and R&D intensity and the R&D types in terms of basic and applied research where input additionality is *more* or *less* likely. The proposed framework also explains why information about *observable* firm or R&D characteristics may not necessarily enable the funder either to achieve optimal subsidy allocations or secure additionality effects when firms invest in basic research and during crisis periods or when firms are older, larger and more R&D-intensive.

3. The subsidy regime as a principal-agent setting

The subsidy regime should be an important starting point in the analysis of effect-size heterogeneity. This is because it sets the rules of the game between the funder and the eligible firms that, in turn, affect: (i) the funder's ability to extract information about the firm's 'true type' in terms of research effort and productivity; and (ii) the R&D investment effort of the subsidised firms after the subsidy. Nevertheless, this is rarely the case in the evaluation literature and the reviews thereof, which tend to adopt a mainly empirical approach that overlooks the issues such as information asymmetries, incomplete contracting, risk aversion or moral hazard. To correct for this oversight, we begin our analysis with a brief discussion of the UK subsidy regime and verify the extent to which it may generate principal-agent problems that affect both subsidy allocations and firm R&D efforts under information asymmetry, incomplete contracting, and risk aversion.

Our sample of firms are eligible to receive direct public support from the government departments, their agencies, and non-departmental public bodies like the *Technology Strategy Board,* including its successor, *Innovate UK*.[4] They are also eligible to receive funding from the European Union (EU) commission.[5]

Despite the involvement of multiple funders, two main features of the subsidy regime stand out. First, the largest part of the subsidies has been managed by non-departmental agencies, of which *Innovate UK* has been the main incumbent until the end of the analysis period. Secondly, R&D subsidies in the United Kingdom (UK) comply with the EU's state-aid rules, under which R&D grants should be related inversely to the so-called market readiness level (MRL) criterion that measures the proximity of the applicant's project to market operations. R&D activities with the lowest score on the MRL scale, for example basic research, are furthest away from the market and qualify for public funding of up to 100% of the project costs. The funding rate gradually declines for R&D activities closer to the market and varies between 25% and 70% of the project cost, depending on firm size (see, Table 1).

To secure public funding, the applicant must demonstrate: (i) the extent to which the project represents value for money for the taxpayer; and (ii) how the applicant will benefit from the innovation, including the latter's impact on productivity and growth. The applicants must also provide convincing information about the scope for additionality, by demonstrating why the proposed project may not be undertaken without public funding.[6] Given this set up, the subsidy regime can be analysed as

Table 1. Funding rates as percentage of eligible project costs: Innovate UK.

Project type→ Firm size↓	Basic research	Feasibility studies	Industrial research	Experimental development
Micro (<10 employees) or Small (<50 employees)	100%	70%	70%	45%
Medium (<250 employees)	100%	60%	60%	35%
Large (250 + employees)	100%	50%	50%	25%

Source: Innovate UK[4].

a principal-agent setting where the principal (the funder) incentivises the agent (the R&D-active firm) to increase its R&D investment to the socially optimal level.

The theory of contracts predicts that the funder can encourage additional R&D investment and maximise social welfare by setting a subsidy that is equal to the 'Pigouvian correction' for positive externalities of the R&D investment. This is a possible outcome if *firms are not risk averse* and *information symmetry exists* between the firm and the funder (Laffont and Martimort 2002; Salanie 2005; Akcigit, Hanley, and Stantcheva 2019). Nevertheless, these necessary conditions for optimal subsidy are rarely, if ever, satisfied in the design of the subsidy regimes.

This seems to be the case in the UK subsidy regime. Figure 1 below depicts the evolution of the subsidy rate (UK R&D subsidies / private R&D intensity) and different R&D input intensities. It indicates that basic R&D intensity has been declining over time and lower than other types of R&D investment. This is the case despite the higher rates of subsidy allowed for basic research in Table 1 and the lack of a downward trend in the overall subsidy rate. Secondly, the data indicates evident declines in all types of R&D intensity just before or during the crisis episodes – the dot-com crisis from 2000 to 2002 and global financial crisis from 2008 to 2010. In the data, the annual basic, private and total R&D intensities are negatively correlated with subsidy rates – with correlation coefficients of −0.1, −0.5, and −0.2, respectively.

The information summarised in Figure 1 is available to the funder at the time of deciding on the level and distribution of the R&D subsidies. This information, however, is insufficient to eliminate the information asymmetry between the funder and the funded firm with respect to four variables that determine the effectiveness of the subsidy regime under the theory of contracts. One is the firm's *risk aversion*, which is private information of the firm and determines the discount rate that the firm utilises to select between R&D projects. The second is the firm's *R&D productivity/profitability*, which is again private knowledge of the firm and determines the firm's optimal level of R&D investment. The third is the firm's *R&D gap*, which reflects the difference between the *firm's optimal R&D investment* and the *socially optimal level* desired by the funder. Finally, the firm has private information about its *market power* (i.e. about its price-cost margin), which the funder does not monitor as s/he takes the existing intellectual property regime as given.

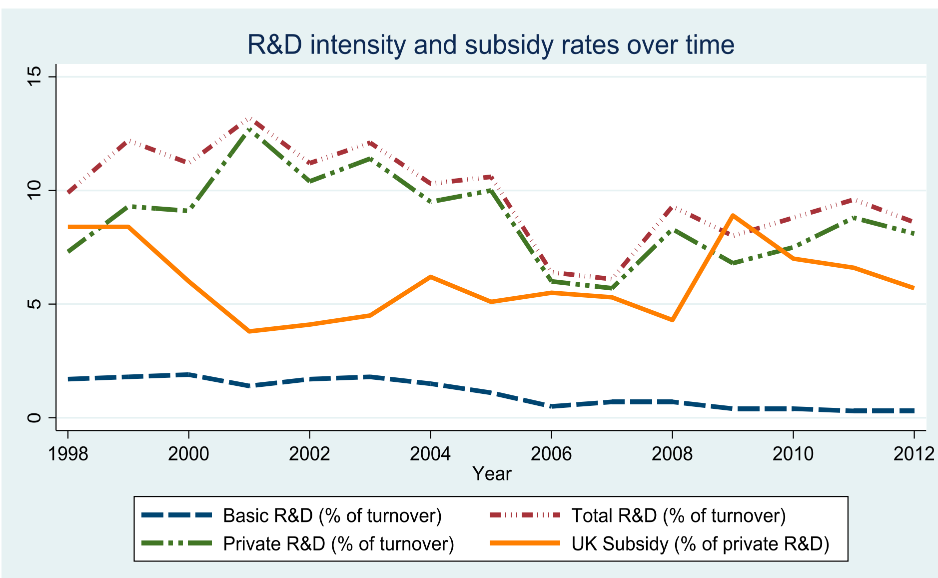

Figure 1. Evolution of R&D intensities and subsidy rates.

Under these conditions, the theory of contracts predicts that subsidy allocations would remain sub-optimal and the additional R&D effort to be induced by the subsidy would be inherently heterogenous (Laffont and Martimort 2002; Salanie 2005; Akcigit, Hanley, and Stantcheva 2019). If firms are heterogenous in terms of their R&D productivity and the latter is private knowledge of the firm, the funder must augment the optimal subsidy – i.e. the 'Pigouvian correction' for R&D externalities – with a 'screening term'. The screening term is necessary to satisfy the participation constraint of the high-R&D-productivity firms and induce the latter to reveal their true types in terms of R&D productivity and hence R&D gap. Moreover, the additional screening term is larger, the higher is the proportion of low-R&D-productivity firms in the applicant pool (Akcigit, Hanley, and Stantcheva 2019).[7]

A second source of deviation from the first-best subsidy is the lack of funder remit to monitor the subsidised firms' pricing and output decisions. The funding agencies take the intellectual property protection regime (hence, the firm's price-cost margin) as given. Therefore, the first-best subsidy is further augmented by a 'monopoly quality valuation correction', which is necessary to induce the firm to produce at the socially optimal level of output (Akcigit, Hanley, and Stantcheva 2019). Yet, and as Antonelli (2020) demonstrates, R&D subsidies are less likely to generate input additionality when the markets for knowledge and/or goods production are not competitive.

A third source of deviation is the firm's risk aversion that, for a given level of R&D productivity, implies a stronger participation constraint for risk-averse firms compared to risk-neutral firms. To satisfy the participation constraint of the risk-averse firms, the principal must offer a larger incentive, which is positively related to (i) the firm's risk-aversion; and (ii) the marginal cost of the extra effort needed to deliver the outcome expected by the funder. Overall, the deviation from the first-best subsidy allocation is larger the higher is the level of risk aversion among the pool of applicants (Laffont and Martimort 2002).

A pertinent question here is whether the funder can achieve first-best outcomes by monitoring the *observable* productivity of the firm's R&D investment – e.g. number of patents, innovative product sales, etc. Demski and Sappington (1984) demonstrate that the principal can achieve first-best outcomes only if the agents' *unobserved* productivities are highly correlated. If the agents are heterogenous, the high-productivity agent can still mimic the low-productivity type and extract informational rents (see also Laffont and Martimort 2002). Moreover, observable proxies of productivity may be ineffective in eliciting effort because they tend to be 'noisy' indicators of true productivity. As such, they cannot be relied upon for contract enforcement (Laffont and Martimort 2002).[8] That is why, the conclusion in Akcigit, Hanley, and Stantcheva (2019) is unequivocal: monitoring the *observed* outcomes of *unobserved* R&D productivity or effort does not enable the principal to elicit truthful information about the agent's true type.

Beyond sub-optimal subsidy allocations, information asymmetry and risk aversion are also sources of heterogeneity in the subsidy's effects on business R&D investment. When firms are heterogenous in terms of R&D productivity/profitability and the correlation between their types is unknown, the high-R&D-productivity firms with narrower R&D gaps can extract informational rents by mimicking the low-productivity firms with wider R&D gaps. Stated differently, the high-R&D-productivity firms with narrower R&D gaps exert less additional effort per unit of subsidy compared to low-R&D-productivity firms with wider R&D gaps.

These insights from the theory of contracts and mechanism design literature indicate that information asymmetry and risk aversion would lead to sub-optimal subsidy allocations and heterogeneity in firm responses to R&D subsidies. Sub-optimal subsidy allocations occur when the funder deviates from the first-best subsidy to satisfy the participation constraint of the high-R&D-productivity firms. On the other hand, heterogenous subsidy effects result from variation in firm types with respect to R&D productivity, risk aversion and the resulting variation in R&D gaps. In the next section, we discuss the extent to which these unobservable firm characteristics can be mapped onto observable characteristics such as age, size, market share and R&D intensity and

R&D types such as basic or applied research that determine the firm's R&D effort hence the scope for input additionality, after the subsidy is allocated.

4. Mapping unobservable information asymmetry and risk aversion onto observable firm characteristics and R&D types

In this section, we draw on a Schumpeterian model of innovation (Aghion, Akcigit, and Howitt 2014, 2015) to relate *observed* firm and R&D types to *unobserved* factors such as R&D productivity, R&D gaps and risk aversion that leads to distortions in subsidy allocations and heterogeneity in the effects of R&D subsidies. In the model, the firm's endogenous R&D intensity increases with the firm's own R&D productivity, which reflects the firm's success in converting R&D investment into innovative product lines. However, R&D intensity decreases with: (a) the rate of creative destruction in the industry that reflects the speed with which technology becomes obsolete; and (b) the firm's discount rate that reflects risk aversion. As such, the firm's *observed* R&D intensity provides vital information about two *unobserved* factors – R&D productivity and risk aversion – that affect the firm's response to the subsidy in the theory of contracts.

Together with unobserved R&D productivity and risk aversion, the firm's observed R&D intensity also provides useful information about observed firm characteristics such as age, size, and proximity to the R&D frontier in the industry. This is because in the Schumpeterian models of innovation, firms survive and grow larger as they invest in R&D and increase the number of innovative product lines. On the other hand, they shrink and eventually exit as their product lines become obsolete (hence, less profitable) due to creative destruction. In this setting, the normalised average value of the innovative product line is stated as follows (Aghion, Akcigit, and Howitt 2014):

$$v_t = \frac{\pi_t \; - \; C_t^{RD}}{\rho_t + x_t - z_{it}} = \frac{\pi_{At}}{\rho_t + x_t - z_{it}} \tag{1}$$

Equation (1) states that value of the innovative product line in year t, v_t, is increasing with adjusted R&D profitability (π_{At}) in the numerator. The adjusted R&D profitability is equal to difference between gross profits, π_t and the cost of R&D investment, C_t^{RD}; and reflects the firm's R&D productivity. The value of the innovative product line is also increasing with R&D intensity, z_{it}, but decreasing with the creative destruction rate, x_t, and the discount rate, ρ_t, in the denominator.

Secondly, the firm's market value, V_t, is given in (2). This is an increasing function of three parameters: (i) the number of innovative product lines (k_t); the normalised average value (v_t) of each innovative product line; and the output per product line (Y_t).

$$V_t(k) = k_t v_t Y_t \tag{2}$$

The firm's market value in (2) is an indicator of observed firm size and its market share. This is because it embodies both the output per product line and the value of that output. Hence, the implication from equation (2) is that firms are larger and/or have larger market shares because their R&D productivity/intensity, the average profitability of their innovative product lines, and the output per product line are higher.

The third prediction relates to the firm's endogenous R&D intensity, z_{it}, which is chosen to maximise the contributions of the innovative product lines to the firm's market value.

$$z_{it} = \left(\frac{v_t}{Wage\_cost_t}\right)^{(1/\eta - 1)} = \left(\frac{(\pi_{At}/\rho_t + x_t - z_{it})}{Wage\_cost_t}\right)^{(1/\eta - 1)} \tag{3}$$

In (3), $Wage\_cost_t$ is the cost of employing R&D scientists and technicians; $1/\eta$ is the elasticity of innovation with respect to employment of scientists and technicians; and the remaining variables are as defined above. It can be observed that R&D intensity, z_{it}, is increasing with R&D productivity (adjusted profitability of the innovative product line); but it is decreasing with the discount rate (ρ_t), the rate of creative destruction (x_t) and wage cost.

The Schumpeterian model also allows for tractable predictions about the relationship between firm survival (hence, firm age) and firm characteristics in terms of R&D intensity and R&D profitability. If the firm's market value follows a Wiener process until liquidation (McDonald and Siegel, 1985), Ugur, Trushin, and Solomon (2016a) relates survival time to R&D intensity/profitability and the discount rate as indicated below.

$$E[t] \cong \frac{2}{2\mu - \sigma^2}\left[\ln(k_t) + \ln\left(\frac{Y_t}{V_0}\right) + \ln\frac{\pi_{At}}{\rho_t + x_t - z_{it}}\right] \tag{4}$$

Here $E[t]$ is expected time until exit, μ and σ drift and volatility parameters of the firm's market value, V_0 is the firm's initial market value, and the rest as defined above. One implication from (4) is that firm's survival time (hence its observed age) is a positive function of R&D productivity/profitability (π_{At}) and R&D intensity (z_{it}).

Given these predictions from the Schumpeterian model, we now define the firm's R&D gap in year t as the difference between the *socially optimal* R&D intensity ($\bar{Z}_t$) and the firm's *privately optimal* R&D intensity (z_{it}) implied by equation (3).

$$RD\_gap_{it} = \bar{Z}_t - z_{it} = \bar{Z}_t - \left(\frac{v_t}{Wage\_cost_t}\right)^{(1/\eta - 1)} = \bar{Z}_t - \left(\frac{(\pi_{At}/\rho_t + x_t - z_{it})}{Wages\_cost_t}\right)^{(1/\eta - 1)} \tag{5}$$

One implication from equation (5) is that the R&D gap ($\bar{Z}_t - z_{it}$) is *narrower* the higher is the R&D profitability/productivity (π_{At}) and R&D intensity (z_{it}). Hence, the effect of the subsidy on the firm's R&D investment would remain small or insignificant if the firm is R&D intensive and has a high level of R&D productivity/profitability. Because firms with higher R&D productivity are also larger (equations 1 and 3), more R&D-intensive (equation 3), and older (equation 5), the case for R&D subsidies is weaker and the probability of input additionality is lower among old, large, and R&D-intensive firms. These outcomes from the Schumpeterian model of innovation dovetail with the finding from the mechanism design model of Akcigit, Hanley, and Stantcheva (2019), where firms with higher R&D productivity (hence with narrower R&D gaps) extracts information rents – i.e. they exert less effort per unit of subsidy. These firm types are able to extract informational rents by concealing their true R&D gaps and mimicking the low-R&D-productivity firms with wider R&D gaps.

The second implication is that the R&D gap is wider when the discount rate (ρ_t) increases due, for example, a perceived increase in the level of return uncertainty associated with R&D investment.[9] The wider R&D gap due to higher levels of return uncertainty strengthens the case for R&D subsidies – in line with the financing constraints argument (Minton and Schrand 1999; Hall 2002, 2010; Czarnitzki and Toole 2007). Indeed, the subsidy increases the firm's incentive to invest in R&D by reducing the R&D cost (C_t^{RD}) and increasing the R&D profitability (π_{At}) in equation (1). Nevertheless, the increase in R&D investment due to the subsidy will be dampened if the firm's discount rate increases at the same time. Hence, the implication of higher return uncertainty and higher risk aversion can be stated as follows: the welfare case for R&D subsidies is stronger but the R&D subsidy may be less effective in generating additionality effects when firms are risk averse and R&D investments are associated with higher levels of return uncertainty.

In what follows, we draw on these insights to derive testable hypotheses on why the effects of R&D subsidies would differ by *observable* firm characteristics (e.g. age, size, and proximity to the R&D frontier) and by return uncertainty that differs between basic and applied research or between crisis and non-crisis periods. In doing this, we begin with the stylised fact that the R&D subsidy, *ceteris paribus*, reduces the R&D cost, increases R&D profitability, and induces the firm to increase its R&D intensity (equation 3). However, once the ceteris paribus assumption is relaxed, equation (3) also indicates that the increase in R&D investment due to the subsidy will be dampened if the discount rate also increases. Such an increase in the discount rate is possible when investment is in basic research with higher return uncertainty; or when the onset of an economic crisis increases the return uncertainty across all types of R&D investment. Under these circumstances, the increase in the discount rates counter-balances the positive effect of the subsidy on the firm's optimal R&D intensity in equation (3). This mitigation effect leaves the R&D gap in equation (5) wider than what is desirable from the funder's perspective or compared the R&D gap for applied research with lower return uncertainty. This result obtains unless the subsidy increases the profitability of the R&D investment by more than the increase in the discount rate that risk-averse firms introduce in the presence of higher return uncertainty. Therefore, we state our first hypothesis (H1) as follows:

> H1: R&D subsidies are less effective in generating additionality effects when perceived return uncertainty is increases during economic crises or when the firm invests in basic research.

H1 is consistent with empirical findings indicating that R&D investment are less responsive to subsidy when R&D returns are uncertain (Aristei, Sterlacchini, and Venturini 2017; Czarnitzki and Toole 2013; Bloom 2007). It is also in line with other findings indicating that the responses of both subsidised and unsubsidised firms to financial crises are pro-cyclical – i.e. firms tend to reduce R&D investment during downturns in the business cycle (Fabrizio and Tsolmon 2014). Finally, it is also consistent with theoretical findings on increasing returns to 'waiting' when investment return uncertainty is high. In this line of work, firms prefer to defer sunk-cost investments when return uncertainty is high and there are positive returns to waiting (Bernanke 1983; Bloom, Bond, and van Reenen 2007; Bouvatier and Lepetit 2008; Czarnitzki and Toole 2013).

H1 implies that the financial constraint and limited appropriability arguments in favour of subsidising basic research are justified theoretically – particularly when markets fail to fund R&D investments with high return uncertainty (Czarnitzki and Toole 2007; Hall 2002; 2010; Minton and Schrand 1999). However, H1 also allows for risk aversion, which is overlooked in predictions that financing constraints and/or limited appropriability are conducive to higher input additionality after the subsidy. If risk aversion exists and risk-averse firms increase the discount rates in the face of increased return uncertainty, the effect of a given subsidy on the firm's R&D intensity is attenuated in equation (3) and the R&D gap remains larger in equation (5). This is the case irrespective of the enhanced level of R&D subsidies granted to basic research projects or crisis-period R&D projects associated with higher return uncertainty compared to applied R&D or R&D investment during non-crisis periods.

H1 is also consistent with findings from several studies that compare firm investment in basic and applied research. This line of research (Arora, Belenzon, and Patacconi 2018; Becker, Hottenrott, and Mukherjee 2022; Czarnitzki and Toole 2013) demonstrates that the rate of increase in basic research undertaken by large firms has fallen behind the rate of increase in applied (or development) research. This trend has been underpinned by a decline in market value of basic research activities and is compatible with the trend for basic research intensity we observe in our sample (Figure 1 above).[10]

In conclusion, H1 implies that the subsidy may be necessary to correct for capital market failures or knowledge externalities that reduce the levels of R&D investment under higher return uncertainty; but the funders may remain unable to close the R&D gap when risk-averse firms adjust their discount rates upwards in the face of higher return uncertainty.

Our second hypothesis (H2) relates to the firm's proximity to the R&D frontier, measured as the difference between the 95th percentile of the R&D intensity in the industry and the firm's own R&D intensity. The funder observes the firm's proximity to the R&D frontier but cannot observe its true R&D productivity or R&D gap. Given this information asymmetry, the funder satisfies the participation constraint of the firms closer to the R&D frontier by augmenting the subsidy with a 'screening' term to achieve two objectives: (i) inducing the firms closer to the R&D frontier to reveal information about their 'true' R&D productivity and R&D gaps; and (ii) minimising the distortion of the subsidy allocations in favour of low-R&D-productivity firm away from the R&D frontier. Nevertheless, if successful, a subsidised firm closer to the R&D frontier exerts less R&D effort compared to its counterparts further away from the R&D frontier. Therefore, we state our second hypothesis (H2) as follows:

> H2: The effect of the subsidy on the firm's R&D investment is weaker or insignificant the closer is the firm to the R&D frontier in the industry. This is the case irrespective of the R&D investment type.

H2 is consistent with and encompasses diverse findings in the empirical literature on R&D subsidies. For example, Chen, Chen, and Mishra (2020) report that the R&D subsidy creates weak or no additionality effects when firms with high intangible intensity are close to the innovation frontier and competition between them is neck-and-neck. In another study, Lach (2002) observes that funders prefer to fund projects with higher success probability, but such preference reduces the scope for additionality as such projects typically have higher private rates of return and could have been undertaken without a subsidy. Similarly, Wanzenböck, Scherngell, and Fischer (2013) report that R&D-intensive firms are less likely to exhibit behavioural additionality because such firms have the capacity and the experience to undertake the desired R&D projects irrespective of public support.

Our third hypothesis (H3) relates to variation in the R&D gap by firm age, size and market share. Predictions from the Schumpeterian models above indicate that high-R&D-productivity firms are larger, older, and would have larger market shares. Furthermore, such firms would have narrower R&D gaps compared to low-R&D-productivity firms. Because of narrower R&D gaps, larger and older firms, and those with larger market shares, would need to exert less effort per unit of subsidy received. Hence, such firms would be less responsive to R&D subsidy compared to smaller and younger firms and firms with smaller market shares. Therefore, we state H3 as follows:

> H3: The responsiveness of larger and older firms, and those with larger market shares to R&D subsidies is weaker.

The analysis above and the derived hypotheses offer a systematic approach to explaining why effect-size estimates vary between and within studies. As such, they take the debate beyond the partial and eclectic explanations that Zúñiga-Vicente et al. (2014) have identified. In what follows, we first discuss our empirical strategy. Then we provide evidence on how the effect of the subsidy on business R&D investment varies by R&D and firm types by the phases of the business cycle.

5. Data and empirical strategy

5.1. Data

Our data is from the Business Research and Development Database (BERD) and Business Structure Database (BSD).[11] The BERD survey covers R&D-active firms stratified by product group and employment size-bands. The reporting units are asked to state intramural (in-house) and extramural (contracted-out) R&D expenditures. For intramural R&D, they are also asked to provide a breakdown by current and capital R&D expenditure. Moreover, the current R&D expenditure is broken down as basic research ('work undertaken for the advancement of scientific knowledge without a specific practical application in view'), applied research ('research undertaken with a general or particular application in view') and experimental research ('results of the basic and applied research directed to the introduction of new materials, processes, products, devices and systems … ').[12] Finally, the firms are asked to state the intramural R&D is financed privately by the firm itself, from UK public funds or from the European Union (EU) funds.[13]

In the dataset, the firm is treated (i.e. it receives UK R&D subsidy) in year *t* if it has been successful in its application beforehand and *has incurred R&D expenditures in that year* to implement the funded project(s). Furthermore, the dataset provides information about total subsidy received from UK sources and disaggregated information about the types of R&D investment undertaken, including total R&D, privately funded R&D, applied R&D, basic R&D, etc. If the different R&D types are complements, the dataset allows for identifying input additionality with respect to aggregate R&D types such as private R&D as well as sub-components such as basic or applied R&D.

We merge the BERD with the Business Structure Database (BSD), which is the universe of all UK firms registered for value-added tax (VAT) and on the pay-as-you-earn (PAYE) system. The BSD contains information on firm turnover, employment, age, survival status, etc. – which we utilise to classify the firms by age, size, and R&D intensity. The merged datasets yield a sample of 43,650 firms observed from 1998 to 2012. Table 2 provides descriptive information about subsidy allocations and subsidy rates by the recipient firm's age and size (employment) deciles.

The share of the subsidy allocated to firms in the top 50% of the distribution is disproportionately higher, at 85.3% and 98.4%, respectively. If the evidence lends support to our third hypothesis (H3), these patterns suggest that the subsidy allocations tend to favour firms that are less likely to create additionality.

Table 2. Private R&D expenditures and UK subsidies by age and size deciles.

	Private R&D	Private R&D intensity	Subsidy	Subsidy rate	Coverage
Panel A – By age deciles	(£ bn.)	(Private R&D as % of turnover)	(£ bn.)	(Subsidy as % of private R&D)	(share of subsidised firm-years, %)
1st decile: age ≤ 3 years	1.27	4.2	0.14	11	96
2nd decile: 3 < age ≤ 6 yrs.	3.25	3.8	0.14	4	94
3rd decile: 6 < age ≤ 9 yrs.	6.57	3.4	0.77	12	93
4th decile: 9 < age ≤ 11 yrs.	8.46	4.6	0.54	6	93
5th decile: 11 < age ≤ 14 yrs.	14.50	4.1	0.57	4	93
6th decile: 14 < age ≤ 17 yrs.	15.20	3.3	0.95	6	92
7th decile: 17 < age ≤ 22 yrs.	29.10	3.3	2.26	8	92
8th decile: 22 < age ≤ 26 yrs.	26.00	2.3	2.85	11	90
9th decile: 26 < age ≤ 31 yrs.	31.20	2.4	3.03	10	91
10th decile: age > 31 years	59.40	2.0	3.43	6	90
Share of top 50%	82.5%		85.3%		
Panel B – By size deciles					
1st decile: 1 employee	0.23	1.5	0.03	14	96
2nd decile: 2 employees	0.25	6.1	0.03	12	97
3rd decile: 3 or 4 employees	0.31	3.6	0.04	12	96
4th decile: 4 < employees ≤ 9	0.70	2.8	0.07	10	95
5th decile: 9 < employees≤ 15	0.95	1.7	0.06	7	94
6th decile: 15 < employees≤ 25	1.52	2.9	0.09	6	94
7th decile: 25 < employees≤ 43	2.49	2.3	0.13	5	93
8th decile: 43 < employees≤ 83	4.93	2.0	0.22	4	92
9th decile: 83 < employees≤ 205	11.20	2.4	0.34	3	91
10th decile: >205 employees	172.00	2.6	13.70	8	80
Share of top 50%	98.7%		98.4%		

Note: Excludes firm/year observations with private R&D intensity greater than 1.
Pooled panel of 43,650 firms with 154,980 firm/year observations.

Table 3. R&D intensity and firm characteristics by treatment status.

	Non-subsidised (Untreated)		Subsidised (Treated)		Whole sample	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
1. Public subsidy (£1,000)	0	0	79.09	3000.25	73.02	2882.97
2. Privately funded R&D (£1,000)	6065.9	45717.4	545.90	15372.05	969.28	19509.76
3. Private R&D intensity	.058	.146	.089	.150	.087	.150
4. Total R&D intensity	.058	.147	.101	.178	.098	.176
5. UK subsidy funding rate	0	0	.009	.042	.009	.040
6. R&D personnel employed	34.765	151.847	5.559	83.896	7.835	91.254
7. R&D personnel intensity	0.080	.236	.095	.194	.094	.198
8. Firm age (years)	19.244	10.263	17.087	10.386	17.253	10.392
9. Deflated turnover (£1,000)	153105	1081527	28371.2	425118.5	37938.2	507617.7
10. Firm employment	70.952	8.125	20.863	5.680	22.897	6.025
11. Start-up dummy	.133	.339	.256	.436	.247	.431
12. Young firm dummy (< 7 years old)	.154	.361	.213	.409	.208	.406
13. Mature firm dummy (> 14 years old)	.624	.484	.537	.498	.544	.498
14. Old firm dummy (> 24 years old)	.353	.478	.264	.441	.271	.444
15. Small firm dummy (< = 25 employees)	.338	.473	.565	.496	.547	.498
16. SMEs (50–250 employees)	.259	.438	.225	.418	.228	419
17. Large firm dummy (> 250 employees)	.283	.450	.085	.279	.101	.301
18. Survivor firm dummy for the whole time	.707	.455	.744	.435	.742	.437
19. Extramural R&D intensity	.006	0.034	.006	.021	.006	.022
20. Capital R&D expenditures intensity	.004	0.015	.005	.011	.005	.011
21. R&D tax credit dummy, SMEs 2008+	.145	.352	.393	.488	.372	.483
22. R&D tax credit dummy, large firms 2008+	.048	.215	.011	.106	.014	.117
23. R&D tax credit dummy, SMEs 2000+	.587	0.492	.850	.357	.829	0.376
24. R&D tax credit dummy, large firms 2002+	.225	0.418	.066	.249	.079	0.270
25. Interaction dummy: SMEs subsidy*R&D tax	.266	.442	.795	.403	.746	.435
Observations	10282		133563		143845	

Notes: + indicates that the R&D tax credit policy changes started from that year. Minimum and maximum values are excluded to comply with non-disclosure requirements of the data owners. Excludes firm/year observations with private R&D intensity greater than 1. The number of firm-year observations may differ from those reported in other tables due to sample-specific one-year-lagged covariates.

Key characteristics of the subsidised and non-subsidised samples are summarised in Table 3.[14] Compared to non-subsidised counterparts, subsidised firms spend less on R&D (row 2) and employ less R&D personnel (row 6). The subsidised firms are also smaller than the non-subsidised firms in terms of turnover (row 9) and total employment (row 10). Nevertheless, in terms of the R&D input intensity, subsidised firms have a relatively higher R&D intensity (rows 3 and 4) and higher R&D personnel intensity (row 7) compared to non-subsidised firms.

The proportion of start-ups (row 11) and young firms (row 12) in the subsidised sample is higher than their proportion in the non-subsidised sample. In contrast, the proportion of mature firms (row 13) and old firms (row 14) is higher in the non-subsidised sample. However, the difference between the proportions of old and mature firms in the subsidised and non-subsidised samples is smaller than the comparable difference for start-ups and young firms. We observe a similar pattern with respect to SMEs (row 16). Finally, surviving firms have similar proportions in both subsidised and non-subsidised samples.

5.2. Empirical strategy

Our aim is to estimate the effect of public funding on various types of business R&D inputs, including privately funded R&D, applied R&D, experimental R&D, basic R&D, extramural R&D, R&D capital expenditures, and employment of R&D personnel. In the data, the firm is treated if it receives UK R&D subsidy for *any type* of R&D investment indicated above. As such, the data *does not* allow for pairing a specific subsidy with a specific R&D type. This is a common data constraint that affects most of the literature on R&D subsidies or R&D tax credits. The problem is addressed by assuming (often implicitly) that different types of R&D expenditures are complements. Such assumptions are justified by existing research, which indicates complementarity between different R&D types (see, for example, Lokshin, Belderbos, and Carree 2008; Mohnen, Polder, and van Leeuwen 2018; Mantovani 2006).

Hence, we assume that a subsidy in any year will affect the aggregate measures and the sub-components of the R&D investment in the year or thereafter in the same direction. A similar identification/estimation strategy has been adopted in past research, of which Aerts and Thorwarth (2008) and Czarnitzki, Hottenrott, and Thorwarth (2011) are two examples.

Our effect-size estimator is the average treatment effect on the treated (ATT) – the conditional difference in mean outcomes for treated (subsidised) and untreated (non-subsidised) firms in the sample. We use two outcome measures: (i) the logarithm of the R&D input's ratio to turnover and the share of scientists and technicians in total employment; and (ii) the annual growth rate (i.e. the first difference) of these outcome measures. Our preferred measure is the latter because first-differencing eliminates the firm-specific fixed effects.[15]

The ATT estimator compares a firm's R&D input intensity when it receives the subsidy with the same firm's unobserved counterfactual outcome in the absence of the subsidy. Given that the firm cannot be observed in both states at the same time, we construct a counterfactual by selecting control (untreated) firms that are close to the treated firms with respect to a range of pre-treatment characteristics (covariates) that affect both selection into treatment and the treatment outcome (Heckman, Ichimura, and Todd 1998; Smith and Todd 2005).

A variety of pre-processing methods exist for constructing a control group from observational data. One approach is to ensure comparability between the treated and untreated samples by estimating propensity scores and using the latter to match the treated and control groups. Then, the expected value of the sample average of the treatment effect on the treated (*SATT*) can be stated as follows:

$$E[SATT] = E[Y_{\boldsymbol{it}}(1)|\ X_{\boldsymbol{it}-1}, D_{it} = 1] - E[Y_{\boldsymbol{it}}(0)|\ X_{\boldsymbol{it}-1}, D_{it} = 1] \tag{6}$$

Here, i and t index firm and year; E is the expectation operator; D_{it} is the binary indicator that is 1 if a firm receives R&D subsidy and 0 otherwise; $Y_{\boldsymbol{it}}$ is the outcome variable as defined above, with $Y_{\boldsymbol{it}}(1)$ indicating the outcome for the treated firm and $Y_{\boldsymbol{it}}(0)$ indicating the outcome for untreated firm. We use one-year-lagged covariates ($X_{\boldsymbol{it}-1}$) covariates that may affect selection into treatment and the measured outcome. The use of lagged covariates reduces the risk of simultaneity in the covariate balancing model.

The conditional expectation $E[Y_{\boldsymbol{it}}(1)|\ X_{\boldsymbol{it}-1}, D_{it} = 1]$ can be estimated directly from the observed sample of treated firms, but the conditional expectation $E[Y_{\boldsymbol{it}}(0)|\ X_{\boldsymbol{it}-1}, D_{it} = 1]$ is the unobserved counterfactual. The latter is estimated by using a control group of firms that are equivalent to the treated firms and would be eligible for treatment with certain probabilities (propensity scores). Using the propensity scores (p_{it}) to construct propensity weights (w^p_{it}), the counterfactual outcome can be estimated as follows:

$$E[Y_{it}(0)|D = 1] = \frac{\sum_{\{it|D=0\}} Y_{it} w^p_{it}}{\sum_{\{it|D=0\}} w^p_{it}} \quad \text{where } w^p_{it} = \frac{p_{it}}{1 - p_{it}} \tag{7}$$

The propensity weights will ensure that the control group is balanced with the treated group if the propensity scores are correct (Hirano, Imbens, and Ridder 2003). However, true propensity scores are unknown, and their estimates are model-dependent. Moreover, it may be difficult to balance all pre-treatment covariates jointly when the selection into treatment is complex and/or the data is highly multidimensional. Finally, the propensity score methods ensure covariate balance only asymptotically – even if the propensity score scores are estimated correctly (Hainmueller 2012; Hainmueller and Xu 2013).

To overcome these limitations, Hainmueller (2012) proposes to estimate the counterfactual outcome with entropy balance (EB) weights, w^{eb}. The EB weights are chosen through a weighting scheme that minimises an entropy distance metric, which decreases with the base weight. The EB weights are then used to obtain the population average treatment effect on the treated (PATT) using the sample estimate in accordance with (8).

$$E[SATT] = E[Y_{\boldsymbol{it}}(1)|\ X_{\boldsymbol{it}-1}, D_{it} = 1] - \frac{\sum_{\{it|D=0\}} Y_{it} w_{it}^{eb}}{\sum_{\{it|D=0\}} w_{it}^{eb}} \tag{8}$$

The EB weights can be obtained for up to three sample moments – i.e. mean, variance, and skewness – of the observables covariates we aim to balance between the treated and control groups. These weights minimise the dissimilarity between the probability distributions of the control and treated groups (Kullback 1959). We derive the EB weights by utilising the Stata routine developed by Hainmueller and Xu (2013). We obtain covariate balancing for a total of 139 covariates (see, below and Table A1 in the Appendix). Then, we estimate the ATT with weighted least squares (WLS), using EB weights as analytical weights.

We account for sources of heterogeneity in the treatment effect in two ways. Our preferred method is to estimate ATTs with different samples. First, we estimate ATTs for different R&D types (private R&D, applied R&D, basic R&D, etc.) over the full sample, Secondly, we estimate ATTs for two crisis periods, one for the *dot.com* bubble burst during 2000–2002; and one for the global financial crisis from 2008 to 2010. Third, we estimate ATTs using four quartiles of the distribution for firm age, size, market share, and R&D intensity. Finally, we estimate ATTs for firms in the manufacturing sector only and those that survive throughout the entire period. Our second approach to sources of heterogeneity is to regress the estimated ATTs for the full sample on firm type, R&D type and crisis period variables in line with Czarnitzki and Lopes-Bento (2013). Congruence between the findings from both methods can be interpreted as an indicator of robustness.

The EB methodology offers several advantages compared to other methods such as matching or propensity score weighting; and produces smaller bias than conventional doubly robust estimators (Zhao and Percival 2017; Amusa, Zewotir, and North 2019).[16] However, the method poses some challenges when the treated and control groups have different sizes or the variance of the weights is too large. We address these issues and other issues that may arise from the panel structure of our data by conducting a battery or robustness checks discussed and evidenced in the online Appendix. These robustness checks include: (i) bootstrapping of the standard errors to take account of potential time-series dependence as recommended by Wooldridge (2010); (ii) introducing several indicator variables to capture multiple subsidies and changes in the R&D tax credits system; (iii) comparing the ATT estimates from the EB methodology with difference-in-difference (D-i-D) estimates that are purged of time-invariant unobserved heterogeneity; and (iv) taking account of omitted unobservable factors in accordance with Oster (2019). The findings from these robustness checks are reported in Tables A7–A9 in the on-line Appendix and indicate a high degree of consistency with the estimates from the EB methodology.

6. Results

The summary statistics for pre-treatment covariates in the EB weighting model are presented in Table A1 in the on-line Appendix. They include 19 covariates at the firm level; 7 covariates at the industry level, including Pavitt technology classes (Pavitt 1984); 6 dummy variables that capture the receipt of EU subsidies and the change in the UK's R&D tax regime; and 2 crisis dummies for 2000–2002 and for 2008–2010. The EB weighting model also include 15 year-dummies and 84 SIC

2-digit industry dummies, but these are not reported in the table to save space. All firm-and industry-level covariates are lagged one year so that treated and untreated firms are balanced one year before treatment in year *t*.

Balancing quality evidence is reported in Tables A11–A15 in the on-line Appendix for a selection of samples.[17] As a general rule, we sought covariate balance for two moments – mean and variance. We measure the quality of covariate balancing with the standardised mean difference between the covariates for treated and *reweighted* control groups. The results are reported in the last two columns of Tables A11–A15.[18] The standardised mean differences after EB weighting (column 11 in bold) are zero or practically zero. The EB weights ensure a satisfactory level of covariate balancing with respect to the second moment (the variance) too – as can be seen by comparing the pre – and post-EB variances in columns 5 and 6. Overall, the improvement covariate balancing reduces the risk of biased ATT estimates.

Table 4 reports ATT estimates for different R&D input types over the full estimation period from 1998 to 2012 (column 1) and during two subperiods that correspond to the *dot.com* crisis (column 2) and global financial crisis (column 3). As indicated above, the ATT is based on the growth rate of the R&D input intensities we analyze.[19] In the case of privately funded R&D, the ATT indicates additionality if it is positive and significant; crowding-out if it is negative and significant; and no effect if the ATT is insignificant. Other R&D inputs such as basic R&D, experimental R&D, or employment of R&D personnel would include both private R&D investment and the public subsidy. Therefore, a positive and significant ATT estimate for these R&D input types would indicate either *additionality* or *absence of full crowding-out* effects, whereas an insignificant or negative ATT indicates *crowding-out* effect.

The results in Table 4 indicate that investments in different R&D types are complements. This is because input additionality over the full sample (column 1) is observed with respect to both aggregate measures such as privately funded R&D and employment of R&D personnel (rows 1 and 2) and disaggregated R&D types such as applied, basic, capital, or extramural R&D investments in the remaining rows. Although the ATTs for disaggregated measures are relatively small, they are significant and indicate complementarity. This finding indicates that, on average, R&D subsidies induce firms to increase both aggregate R&D investment and its sub-components. This is consistent with evidence in the wider literature on complementarity between different R&D types (Lokshin, Belderbos, and Carree 2008; Mantovani 2006; Mohnen, Polder, and van Leeuwen 2018).

However, the ATT for basic R&D (row 5 of column 1) is smaller than the ATT for private R&D (row 1). The basic R&D intensity of subsidised firms increases by 0.6% over the analysis period compared to a much higher growth rate of 4.6% in the case of private R&D or R&D personnel intensity. We interpret this finding as evidence that lends support to our first hypothesis (H1), which postulates that R&D subsidies are less effective in generating additionality effects on basic R&D due to lower market readiness and higher return uncertainty associated with basic research. Further evidence that also lends support to H1 is reported in columns 2 and 3 of Table 4, where the ATTs are

Table 4. Effects of R&D subsidy on different R&D types: comparing full analysis period with crisis periods.

Subsidy's effect on growth rate of:	(1) Full sample	(2) dot-com crisis 2000–2002	(3) Global financial crisis 2008–2010
Private R&D intensity	.0457*** (.0060)	.0217 (.0167)	.0235*** (.0075)
R&D personnel intensity	.0456*** (.0066)	.0129 (.0151)	.0365*** (.0111)
Applied R&D intensity	.0244*** (.0036)	.0153* (.0092)	.0077 (.0079)
Experimental R&D intensity	.0158*** (.0072)	.0011 (.0104)	.0214*** (.0052)
Basic R&D intensity	.0063*** (.0015)	.0113*** (.0040)	.0019*** (.0005)
Extramural R&D intensity	.0027*** (.0006)	.0014 (.0021)	.0011 (.0015)
Capital R&D intensity	.0040*** (.0007)	.0035* (.0018)	.0024** (.0009)
Observations in control sample	$N_0 = 10282$	$N_0 = 1821$	$N_0 = 3510$
Observations in treated sample	$N_1 = 133563$	$N_1 = 15955$	$N_1 = 38934$

Note: The outcome variable is the growth of R&D input intensity. Bootstrap standard errors are estimated with 100 iterations. N_0 is number of firm-year observations in the control (unsubsidised) sample; N_1 is number of firm-year observations in the treated (subsidised) sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

usually smaller in magnitude more likely to remain statistically insignificant during the crisis periods in our data. This is the case for most R&D types – except for experimental R&D during the global financial crisis (row 4) and basic R&D during the *dot.com* crisis (row 5). These findings are also consistent with R&D investment models under uncertainty (Czarnitzki and Toole 2013) and 'waiting' models of investment (Bloom 2007), where R&D investment is less responsive to policy interventions when return uncertainty is higher. The findings also provide a theory-informed explanation for empirical findings in Aristei, Sterlacchini, and Venturini (2017), who report weaker additionality effects during the global financial crisis.

The evidence in Table 4 begs the question of whether large or small firms are responsible for the small input additionality in basic research. The findings in Table 5 provide a clear answer: input additionality in basic research is generated by small and young firms and those with smaller market shares (quartiles 1 and 2). These findings are consistent with trends reported in Arora, Belenzon, and Patacconi (2018) and Becker, Hottenrott, and Mukherjee (2022), who report higher rates of basic R&D investment among small firms as opposed to large firms. They are also in line with our data, where basic R&D intensity is indeed higher among small and young firms in the bottom half of the age and size distributions.

Yet, the evidence in Tables 4 and 5 and further evidence below indicate another trend too: input additionality in the case of basic R&D is always smaller than private or applied R&D intensity irrespective of the firm size. Hence, we conclude that the responsiveness of *basic R&D intensity* to the UK subsidy has been weaker irrespective of the firm types and whether the analysis is over the crisis period (Table 4) or over the full period (Table 5).

The findings in Tables 4 and 5 enable us to revisit the case for higher levels of public support when the firm's financing constraint is stronger due to return uncertainty (Hall 1992, 2002, 2010; Minton and Schrand 1999; Czarnitzki and Toole 2007). Higher levels of support are justified under these conditions because they reduce the R&D cost and may enable the financially constrained firms to reduce their R&D gaps, leading to improvement in social welfare. However, both equation (5) from the Schumpeterian model of innovation and findings in Tables 4 and 5 indicate that the effect of the subsidy on the R&D gap is attenuated when the investment is in basic research or undertaken during crisis episodes, which induce the risk averse firms to increase the discount rates that they use for selecting between R&D investment projects. Hence, under the conditions of return uncertainty and risk aversion, subsidies may be necessary to alleviate the 'financing constraint' of the firms investing in basic research or during crisis episodes, but the input additionality under these conditions is small and may be practically insignificant.

Noting that funders subsidise up to 100% of the project cost for basic research and the subsidy intensity or coverage rates are higher during crisis periods, we identify the first *conundrum* in public

Table 5. Effect of R&D subsidies on growth of basic R&D intensity: estimates by firm-type quartiles.

	Basic R&D by age quartile	Basic R&D by size quartile	Basic R&D by market-share quartile
Quartile 1	.0067** (.0033)	.0148*** (.0042)	.0177*** (.0067)
	N0 = 2189	N0 = 1081	N0 = 831
	N1 = 33154	N1 = 35765	N1 = 35103
Quartile 2	.0062** (.0028)	.0038*** (.0019)	.0051*** (.0016)
	$N_0 =$ 1948	$N_0 =$ 1781	$N_0 =$ 1803
	$N_1 =$ 33668	$N_1 =$ 33303	$N_1 =$ 34175
Quartile 3	.0044 (.0024)	.0002 (.0006)	.0017* (.0009)
	$N_0 =$ 2325	$N_0 =$ 2381	$N_0 =$ 2751
	$N_1 =$ 33054	$N_1 =$ 33562	$N_1 =$ 33401
Quartile 4	.0006 (.0004)	.0002 (.0002)	.0015 (.0015)
	$N_0 =$ 2971	$N_0 =$ 5028	$N_0 =$ 5006
	$N_1 =$ *32683*	$N_1 =$ *30921*	$N_1 =$ *30396*

Notes: Age is measured in years; size is measured by number of employees; market share is the share in industry turnover at 3-digit SIC industry level (254 industries). The sample size for different R&D types differs because firms do undertake all types of R&D investment every year. For other notes, see Table 4 above.

support for business R&D as follows: *it may be socially desirable to allocate higher levels of subsidy for basic research or during economic downturns, but business R&D would be less responsive to the subsidy under these conditions*.

Table 6 presents the results for the subsidy's effects on the growth of different R&D input intensities, depending on the firms' distance to the R&D frontier in the 3-digit industry. The results indicate clearly that the ATTs are *insignificant* in the first and second quartiles of the distance to the R&D frontier, where firms are closer to the frontier defined as the 95th percentile of the relevant R&D input intensity in the industry and year. The effect-size becomes significant but remains small in quartile 3; and it is relatively larger in quartile 4, where firms are the furthest away from the R&D frontier. This is the case irrespective of the R&D input.

These findings are in line with H2, which stipulates that the subsidy's effect on various types of business R&D investment is smaller the closer is the firm to the R&D frontier in the industry. This is because R&D intensity among firms closer to the R&D frontier are higher and their R&D gaps are smaller even in the absence of the subsidy, as implied by equation (2) of the Schumpeterian innovation model. Given these narrower R&D gaps, the subsidy would induce smaller additional R&D investment as the firms are already close to the R&D frontier, which represents the highest level of privately optimal R&D intensity in the industry.

The evidence in Table 6 reveals a second *conundrum* in public support for business R&D:

> it is socially desirable to support high-R&D-productivity firms closer to the R&D frontier as they are more likely to be successful in converting R&D inputs into successful innovations, but firms closer to the R&D frontier are less likely to generate additionality effects as their R&D gaps are narrower.

This conundrum is consistent with theoretical insights from the theory of contracts (Laffont and Martimort 2002) and mechanism design models (Akcigit, Hanley, and Stantcheva 2019), where high-R&D-productivity firms have lower R&D gaps but are more likely to extract informational rents by concealing their true types. It is also consistent with empirical findings in Lach (2002), Lee (2011), and Wanzenböck, Scherngell, and Fischer (2013) who report that subsidies are less likely to generate additionality among firms with high R&D intensity.

Table 6. Effects of R&D subsidies on growth of different R&D input intensities: estimates by R&D type and distance to R&D frontier.

Distance to R&D frontier	Growth of private R&D intensity	Growth of R&D personnel intensity	Growth of experimental R&D intensity	Growth of applied R&D intensity	Growth of basic R&D intensity	Growth of extramural R&D intensity	Growth of capital R&D intensity
Quartile 1	-.0007	.0085	-.0009	.0007	-.0016	-.0017	.0013
N_0 = 2782	(.0084)	(.0077)	(.0070)	(.0029)	(.0025)	(.0023)	(.0009)
N_1 = 22173							
Quartile 2	.0037	.0068	.0027	.0018	.0003	.0004	.0001
N_0 = 1301	(.0053)	(.0051)	(.0042)	(.0023)	(.0005)	(.0009)	(.0004)
N_1 = 23655							
Quartile 3	.0248***	.0249***	.0116***	.0112***	.0018	.0033**	.0038***
N_0 = 485	(.0055)	(.0055)	(.0037)	(.0033)	(.0015)	(.0016)	(.0009)
N_1 = 24470							
Quartile 4	.0495***	.0611***	.0271***	.0263***	.0134***	.0051***	.0076***
N_0 = 470	(.0134)	(.0158)	(.0057)	(.0099)	(.0031)	(.0009)	(.0014)
N_1 = 24486							

Notes: Distance to R&D frontier is measured as the difference between the 95th percentile of the total R&D intensity in the 3-digit SIC industry (254 industries) and the firm's own R&D intensity. Firms in quartile 1 are closest to R&D frontier. For other notes, see Tables 4 and 5 above.

Table 7. Effects of R&D subsidies by firm age, size, and market share.

Panel A: Effects on growth of privately funded R&D and R&D personnel intensity

	Growth of private R&D intensity by firm			Growth of R&D personnel intensity by firm		
Grouping by →	Age quartile	Size quartile	Market share quartile	Age quartile	Size quartile	Market share quartile
Quartile 1	.0648*** (.0133) N_0 = 2189 N_1 = 33154	.0849*** (.0142) N_0 = 1081 N_1 = 35765	.1120*** (.0210) N_0 = 1046 N_1 = 34682	.0611*** (.0153) N_0 = 2189 N_1 = 33154	.0853*** (.0142) N_0 = 1081 N_1 = 35765	.0909*** (.0124) N_0 = 1046 N_1 = 34682
Quartile 2	.0443*** (.0116) N_0 = 1948 N_1 = 33668	.0275** (.0137) N_0 = 1781 N_1 = 33303	.0333*** (.0095) N_0 = 1990 N_1 = 34018	.0483*** (.0132) N_0 = 1948 N_1 = 33668	.0387*** (.0136) N_0 = 1781 N_1 = 33303	.0447*** (.0116) N_0 = 1990 N_1 = 34018
Quartile 3	.0198 (.0138) N_0 = 2325 N_1 = 33054	.0004 (.0003) N_0 = 2381 N_1 = 33562	.0208 (.0162) N_0 = 2390 N_1 = 33649	.0189* (.0109) N_0 = 2325 N_1 = 33054	.0001 (.0005) N_0 = 2381 N_1 = 33562	.0005 (.0007) N_0 = 2390 N_1 = 33649
Quartile 4	-.0003 (.0006) N_0 = 2971 N_1 = *32683*	.0063 (.0039) N_0 = 5028 N_1 = *30921*	.0006 (.0035) N_0 = 4845 N_1 = *31202*	.0018 (.0011) N_0 = 2971 N_1 = *32683*	.0046 (.0028) N_0 = 5028 N_1 = *30921*	.0004 (.0011) N_0 = 4845 N_1 = *31202*

Panel B: Effects on growth of extramural R&D and capital R&D intensities

	Growth of extramural R&D intensity			Growth of capital R&D expenditures intensity		
Grouping by →	Age quartile	Size quartile	Market share quartile	Age quartile	Size quartile	Market share quartile
Quartile 1	.0038** (.0018) N_0 = 2189 N_1 = 33154	.0047*** (.0016) N_0 = 1081 N_1 = 35765	.0043 (.0052) N_0 = 2189 N_1 = 33154	.0062*** (.0010) N_0 = 2189 N_1 = 33154	.0069*** (.0016) N_0 = 1081 N_1 = 35765	.0108*** (.0017) N_0 = 2189 N_1 = 33154
Quartile 2	.0017 (.0019) N0 = 1948 N1 = 33668	.0002 (.0031) N0 = 1781 N1 = 33303	.0015 (.0023) N0 = 1990 N1 = 34018	.0024 (.0018) N0 = 1948 N1 = 33668	.0043*** (.0009) N0 = 1781 N1 = 33303	.0038*** (.0009) N0 = 1990 N1 = 34018
Quartile 3	.0015 (.0019) N_0 = 2325 N_1 = 33054	.0014 (.0050) N_0 = 2381 N_1 = 33562	.0095 (.0073) N0 = 2390 N1 = 33649	.0016 (.0012) N_0 = 2325 N_1 = 33054	-.0002 (.0006) N_0 = 2381 N_1 = 33562	.0095 (.0073) N0 = 2390 N1 = 33649
Quartile 4	.0008 (.0020) N_0 = 2971 N_1 = *32683*	-.0008 (.0009) N_0 = 5028 N_1 = *30921*	.0008 (.0011) N_0 = 4845 N_1 = *31202*	.0016 (.0011) N_0 = 2971 N_1 = *32683*	-.0008 (.0009) N_0 = 5028 N_1 = *30921*	.0003 (.0003) N_0 = 4845 N_1 = *31202*

Notes: Age is measured in years; size is measured by number of employees; market share is the share in industry turnover at 3-digit SIC industry level (254 industries). Firms in quartile 1 are smallest, youngest, and with smallest market share. For other notes, see Tables 4 and 5 above.

Table 7 presents ATT estimates by firm age, size, and market share for four R&D types. Results in Panel A indicate that the ATTs for privately funded R&D and R&D personnel intensity are declining as firm age, size, and market share increases. The ATTs are largest among youngest and smallest firms and firms with smaller market shares (quartile 1). Then the ATTs decline and become insignificant as firms grow in age or size or capture larger market shares (in quartiles 3 and 4 of the age, size and market share distributions). These patterns clearly indicate the positive effect of the subsidy on private R&D and R&D personnel intensity in (Table 4, column 1) conceals a high degree of heterogeneity. They also indicate that the positive and significant additionality effect in the full sample is driven by the additionality effect among smaller and younger firms. As such, the ATT in the full sample is a poor basis for evidence-based public policy. A similar pattern is evident in panel B, where we report ATTs for extramural R&D and capital R&D investment. In both cases, the ATTs decline and eventually become insignificant as firm age, size or market share increases.

These findings lend support to H3, which stipulates that the subsidy's effect on business R&D would decline as firm age, size, or market value increases. They are also consistent with

Schumpeterian models of innovation (Aghion, Akcigit, and Howitt 2014, 2015), where firms are larger and older due to a history of success in R&D productivity. Finally, they are also consistent with insights from the theory of contracts, where high-R&D-productivity firms are better placed to conceal their true types and extract informational rents under asymmetric information (Laffont and Martimort 2002; Akcigit, Hanley, and Stantcheva 2019).[20]

The results in Table 7 point out to a third conundrum in public support for business R&D:

> Funders tend to allocate subsidies to larger and older firms with proven track records to minimise the distortion of the subsidy allocation in favour less efficient firms, but the firms with proven track record of successful innovations are less likely to produce additionality effects.

This conundrum is particularly acute in the UK, where 85–98% of the subsidy is allocated to firms in quartiles 3 and 4 of the age and size distributions.

We have carried out a wide range of robustness check to verify whether the findings above remain robust to variations in sampling, measurement, and specification choices. The results from the first set of robustness checks are reported in five additional tables in the online Appendix (Tables A2–A6). In Table A2, we checked whether the findings differ when the logarithm of R&D intensity is used as the outcome variable instead of the growth in R&D input intensity. In Table A3, we restricted the sample to surviving firms only to verify if the exclusion of exiting firms leads to different findings. Then we used initial weights from a coarsened exact matching (CEM) procedure (Table A4). In Table A5, we restricted the treated sample to firms that have received UK R&D subsidy only once during the analysis period to verify whether the results in the main text might be driven by multiple treatment (i.e. by receipt of multiple subsidies). Finally, we regressed the estimated ATTs for private R&D intensity growth in the full sample on firm characteristics and crisis periods (Table A6) to verify if the average effect in the sample varies with firm characteristics and crisis episodes. The findings from this first set of robustness checks are consistent with the results reported and discussed above. The ATTs are smaller or insignificant among larger and older firms and firms closer to the R&D fronter. They are also smaller when investment is in basic research or during crisis episodes.

Results from the second set of robustness checks are presented in Tables A7–A9 in the online Appendix. One check consists of estimating the ATTs with a difference-in-difference (D-i-D) methodology, which takes account of the observable factors we control for in the EB weighting model and the unobservables. The other check draws on Oster (2019), who develops a routine that establishes the extent to which the unobservables must be influential to invalidate the significance of the ATTs based on controlling for observables only. The findings from these robustness checks also indicate a high degree of consistency with the estimates from the EB methodology.

In none of the additional robustness checks has the EB-based inference of insignificant ATTs been overturned by significant input additionality inference based on Oster (2019) methodology. The ATTs from EB and Oster (2019) methodologies do differ in Table A7 in 4 out of 16 estimations. The discrepancy, however, is due to higher incidence of insignificant or crowding-out effects when Oster (2019) methodology is used. Moreover, the D-i-D estimates are consistent with EB and Oster (2019) estimates in 70% of the findings on effect-size heterogeneity by R&D type and crisis episodes (Table A7). The level of consistency is even higher at 100% in Table A8, where we report estimates by the firms' proximity to R&D frontier. The level of consistency is 90%–100% in Table A9, where we report effect-size estimates by firm size (employment size) and firm market share.

Therefore, we re-affirm our findings that the effects of R&D subsidies are smaller and tend to be insignificant among larger and older firms and firms closer to the R&D frontier. The effects are also smaller or insignificant when returns on R&D investment are less certain during crisis episodes or when the investment is in basic research.

7. Conclusions

In this paper, we have demonstrated that it is feasible to bridge the gap between theory and empirics with respect effect-size heterogeneity reported in the evaluation literature on R&D subsidies. This is done by drawing on the theory of contracts and Schumpeterian models of innovation, which allow for deriving testable hypotheses on why the subsidy's effect on business R&D may vary by R&D type and firm characteristics. The proposed theoretical framework allows for identifying the sources of heterogeneity in a systematic manner, which we achieve by: (i) unpacking the subsidy contract to demonstrate how unobservable factors such as risk aversion and information asymmetry with respect to unobserved R&D productivity generate heterogeneity in the subsidy's effect on business R&D investment; and (ii) mapping these unobserved factors on to observable factors related to firm characteristics (age, size, market share, distance to R&D frontier) and R&D types (basic, applied, experimental, capital R&D).

We have also expanded the evidence base through findings that are consistent with the proposed hypotheses and remain robust to a wide range of sensitivity checks. The evidence indicates that R&D subsidies are conducive to input additionality in the full sample, but the average input additionality conceals a high degree of heterogeneity. The input additionality effects: (i) are smaller or insignificant when investment is in basic research or during crisis periods; and (ii) decline and eventually become insignificant among larger and older firms and firms closer to the R&D frontier.

Our findings point out to evident conundrums in public support for business R&D. On the one hand, the case for R&D subsidies is stronger when firms invest in basic research, during crisis episodes or have a track record of success in converting R&D investment into profitable product lines. On the other hand, R&D subsidies are less likely to generate additionality effects under these conditions. Therefore, we conclude that R&D subsidies are necessary but not sufficient for maximising social welfare and closing the R&D gap, defined as the difference between actual and socially optimal levels of R&D investment.

Our findings suggest that the case for R&D subsidies as an innovation policy instrument is less clear-cut than the sanguine policy stance in favour of direct public support schemes internationally. They indicate that direct support for business R&D investment should target small and young firms and firms with wider R&D gaps, which are more likely to generate additionality effects compared to large and old firms with narrower R&D gaps. They also indicate that the effectiveness of R&D subsidies in increasing basic research (or all type of R&D investment during crisis periods) is limited.

These qualifications are consistent with several findings in the empirical literature reviewed above. It is also consistent with Antonelli (2020), who demonstrate that small and young firms that operate in competitive markets for knowledge and goods production are more likely to generate input additionality. Finally, the qualified case for R&D subsides we arrive at is also consistent with Akcigit, Hanley, and Stantcheva (2019), who argue for selective support that favours small and young firms, combined with *ex post* 'innovation prizes' that reward successful innovations.

Our findings are based on best practice in the evaluation literature and remain consistent across multiple robustness checks. However, it is necessary to acknowledge two potential shortcomings that may impinge on the quality of the reported evidence. The first relates to the relatively small size of the control group in the data. This data constraint has led us to rely on covariate balancing with respect to the first moment (mean) only in most estimation samples and first and second moment (mean and variance) in a few remaining samples. As a result, residual difference between treated and control group is eliminated with respect to the mean of the observables only. The second relates to the absence of information in the data on the R&D type (e.g. basic, applied, experimental, etc.) for which the R&D subsidy is granted. Because of this constraint, the ATT estimates for disaggregate R&D components are identified by relying on complementarity between R&D investment types reported in the relevant literature.

Notes

1. A recent evaluation study reports that the EB routine is one of the top-5 performers of twenty popular estimators in terms of root means square error (RMSE) and coverage of the true treatment effect, while the propensity score estimators are not (Dorie et al. 2019).
2. It must be indicated at the outset that the focus of this study is on variation in the effects of R&D subsidies and the firm or R&D characteristics that may be conducive to effect-size heterogeneity. Hence, our review of the empirical literature will remain limited to these empirical patterns and whether they can be explained through a unified theoretical framework instead of the eclectic and mostly *ex post* explanations on offer.
3. The existing work indicates demonstrated that knowledge spillovers should not be considered as a 'free lunch'; and that firms must invest in own R&D to exploit knowledge externalities. This line of research indicates that the required level of own R&D investment differs between firms, depending on the firms R&D intensity and the nature of technology in the industry.
4. The non-departmental public agencies also include eight regional development agencies (RDAs), which also provided R&D funding from 2000 to 2012, but then discontinued.
5. We investigate only the effects of UK subsidies on business R&D investment. To ensure identification of the UK subsidy's effect, we balance the treated and control firms by taking account of EU funding in addition to a wide set of other covariates.
6. See Innovate UK, General guidance for grant applicants at https://www.gov.uk/guidance/innovate-uk-funding-general-guidance-for-applicants.
7. The funder agrees to pay the 'screening term' to avoid social welfare losses that would arise from under-representation of the high-R&D-productivity (hence, over-representation of the low-R&D-productivity) firms in the allocations of the R&D subsidies. (Akcigit, Hanley, and Stantcheva 2019).
8. In the best-case scenario limited to one principal and two agents only, a principal that uses observable proxies for productivity can restrict only one of the agents to tell the truth, whilst the other agent adopts a Nash response to the truth (Demski and Sappington 1984).
9. On the relationship between risk aversion, risk premia and discount rates, see Hirshleifer (1961), Chronopoulos, De Reyck, and Siddiqui (2011).
10. In the words of Arora, Belenzon, and Patacconi (2018), large firms " … still value the golden eggs of science (as reflected in patents) but seem to be increasingly unwilling to invest in the golden goose itself (the internal scientific capabilities)".
11. Office for National Statistics (2019a, 2019b). See also Ugur, Trushin, and Solomon (2016b).
12. See examples of the questionnaires at https://www.ons.gov.uk/businessindustryandtrade/business/
13. The privately funded R&D investment we use in this paper is the difference between total intramural R&D expenditures and the sum of UK and EU subsidies.
14. The sample excludes firm-year observation with private R&D intensity greater than 1. The number of excluded firms is 738 with 2,190 firm-year observations. The excluded firms have excessively high levels of private R&D intensity (up to several hundreds) due to very small turnover values. Most of these firms have short survival times and exit during the analysis period (see Ugur, Trushin, and Solomon 2016a).
15. Nonetheless, we conduct sensitivity checks with the logarithm of R&D input intensity.
16. Therefore, the EB method has been considered as a welcome addition to the range of treatment-effect estimators that rely on direct covariate balancing instead of matching (Athey and Imbens 2017). Also, the method has been applied in over 40 evaluation studies in economics and other social science disciplines, examples of which include McMullin and Schonberger (2020), Marcus (2013), and Neuenkirch and Neumeier (2016).
17. Descriptive statistics and entropy balancing information for other samples used in the estimations are not reported here to save space, but are available on request.
18. The two most common metrics for evaluating covariate balance quality are the standardised difference (bias) in means and t-tests for differences in means (Rosenbaum and Rubin 1983). Imai, King, and Stuart (2008) criticise the use of t-tests and argue for QQ plot summary statistics. A recent simulation study (Franklin et al., 2014) reports that standardised difference in means has the advantage of evaluating balance on all covariates simultaneously.
19. We found similar results when the outcome measure is the logarithm of the R&D input intensities instead of growth rates. These are not reported here to save space but are available on request. Furthermore, we find that the subsidy's effect on private R&D intensity is smaller when we regress the effect-size estimate on two crisis dummies (see Table A6 in the on-line Appendix).
20. It must also be noted that the findings in Table 7 are consistent with Lach (2002), González and Pazó (2008), and Wanzenböck, Scherngell, and Fischer (2013), who report that the additionality effect is observed mainly among small firms. They are also in line with the review of Zúñiga-Vicente et al. (2014), who conclude that the subsidy is less likely to induce R&D additionality among large firms.

Acknowledgements

We thank our funders for the generous support. We also thank the data host, the UK Data Service, for their help and support in accessing the data. We also thank Edna Solomon for help with building the dataset. The use of this dataset does not imply the endorsement of the data owner or the UK Data Service in relation to the interpretation or analysis of the data. The views expressed here are those of the authors, who are responsible for any errors or omissions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research has been funded by the ESRC under [grant number ES/K004824/1].

ORCID

Mehmet Ugur http://orcid.org/0000-0003-3891-3641
Eshref Trushin http://orcid.org/0000-0001-9058-4262

References

Office for National Statistics, Business Expenditure on Research and Development, 1995–2017: Secure Access [data collection], 8th Edition, UK data Service. SN: 6690, 2019a. doi:10.5255/UKDA-SN-6690-8.

Office for National Statistics, Business Structure Database, 1997–2018: Secure Access [data collection], 10th Edition, UK data Service. SN: 6697, 2019b. doi:10.5255/UKDA-SN-6697-10.

Aerts, K., and S. Thorwarth. 2008. "Additionality effects of public R&D funding: 'R' versus 'D'.'' FBE Research Report MSI_0811, Faculty of Business and Economics, KU Leuven: 1-19.

Aghion, P., U. Akcigit, and P. Howitt. 2014. "What do we Learn from Schumpeterian Growth Theory?" In *Handbook of Economic Growth, vol. 2*, edited by Philippe Aghion, and Steven N. Durlauf, 515–563. Amsterdam: Elsevier.

Aghion, P., U. Akcigit, and P. Howitt. 2015. "Lessons from Schumpeterian Growth Theory." *American Economic Review* 105 (5): 94–99.

Akcigit, U., D. Hanley, and S. Stantcheva. 2019. "Optimal Taxation and R&D Policies." National Bureau of Economic Research Working Paper No. w22908, Cambridge, MA.

Amusa, L., T. Zewotir, and D. North. 2019. "Examination of Entropy Balancing Technique for Estimating Some Standard Measures of Treatment Effects: A Simulation Study." *Electronic Journal of Applied Statistical Analysis* 12 (2): 491–507.

Antonelli, C. 2020. "Knowledge Exhaustibility Public Support to Business R&D and the Additionality Constraint." *Journal of Technology Transfer* 45 (3): 649–663.

Aristei, D., A. Sterlacchini, and F. Venturini. 2017. "Effectiveness of R&D Subsidies During the Crisis: Firm-Level Evidence Across EU Countries." *Economics of Innovation and New Technology* 26 (6): 554–573.

Arora, A., S. Belenzon, and A. Patacconi. 2018. "The Decline of Science in Corporate R&D." *Strategic Management Journal* 39 (1): 3–32.

Arrow, K. 1996. "The Economics of Information: An Exposition." *Empirica* 23 (2): 119–128.

Athey, S., and G. W. Imbens. 2017. "The State of Applied Econometrics: Causality and Policy Evaluation." *Journal of Economic Perspectives* 31 (2): 3–32.

Becker, B. 2015. "Public R&D Policies and Private R&D Investment: A Survey of the Empirical Evidence."." *Journal of Economic Surveys* 29 (5): 917–942.

Becker, A., H. Hottenrott, and A. Mukherjee. 2022. "Division of Labor in R&D? Firm Size and Specialization in Corporate Research." *Journal of Economic Behavior and Organization* 194 (February): 1–23.

Bernanke, B. 1983. "Irreversibility, Uncertainty, and Cyclical Investment'." *Quarterly Journal of Economics* 98 (1): 85–106.

Bloom, N. 2007. "Uncertainty and the Dynamics of R&D." *American Economic Review* 97 (2): 250–255.

Bloom, N., S. Bond, and J. van Reenen. 2007. "Uncertainty and Investment Dynamics." *Review of Economic Studies* 74 (2): 391–415.

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

Boeing, P., J. Eberle, and A. Howell. 2022. "The Impact of China's R&D Subsidies on R&D Investment, Technological Upgrading and Economic Growth." *Technological Forecasting and Social Change* 174 (January): 1–10.

Bouvatier, V., and L. Lepetit. 2008. "Banks' Procyclical Behavior: Does Provisioning Matter?" *Journal of International Financial Markets, Institutions and Money* 18 (5): 513–526.
Branstetter, L. M. Sakakibara. 1998. "Japanese Research Consortia: A Microeconometric Analysis of Industrial Policy." *Journal of Industrial Economics* 46 (2): 207–233.
Cerulli, G. 2010. "Modelling and Measuring the Effect of Public Subsidies on Business R&D: A Critical Review of the Econometric Literature." *Economic Record* 86 (274): 421–449.
Chen, Y. H., M. X. Chen, and A. K. Mishra. 2020. "Subsidies Under Uncertainty: Modeling of Input-and Output-Oriented Policies." *Economic Modelling* 85: 39–56.
Chronopoulos, M., B. De Reyck, and A. Siddiqui. 2011. "Optimal Investment Under Operational Flexibility, Risk Aversion, and Uncertainty." *European Journal of Operational Research* 213 (1): 221–237.
Clausen, T. 2009. "Do Subsidies Have Positive Impacts on R&D and Innovation Activities at the Firm Level?" *Structural Change and Economic Dynamics* 20 (4): 239–253.
Cohen, W., and D. Levinthal. 1989. "Innovation and Learning: The two Faces of R&D." *Economic Journal* 99 (397): 569–596.
Czarnitzki, D., and H. Hottenrott. 2011. "R&D Investment and Financing Constraints of Small and Medium-Sized Firms." *Small Business Economics* 36 (1): 65–83.
Czarnitzki, D., H. Hottenrott, and S. Thorwarth. 2011. "Industrial Research Versus Development Investment: The Implications of Financial Constraints." *Cambridge Journal of Economics* 35 (3): 527–544.
Czarnitzki, D., and C. Lopes-Bento. 2013. "Value for Money? New Microeconometric Evidence on Public R&D Grants in Flanders." *Research Policy* 42 (1): 76–89.
Czarnitzki, D., and A. Toole. 2007. "Business R&D and the Interplay of R&D Subsidies and Product Market Uncertainty." *Review of Industrial Organization* 31 (3): 169–181.
Czarnitzki, D., and A. Toole. 2013. "The R&D Investment–Uncertainty Relationship: Do Strategic Rivalry and Firm Size Matter?" *Managerial and Decision Economics* 34 (1): 15–28.
David, P., B. Hall, and A. Toole. 2000. "Is Public R&D a Complement or Substitute for Private R&D? A Review of the Econometric Evidence." *Research Policy* 29 (4-5): 497–529.
Demski, J. S., and D. Sappington. 1984. "Optimal Incentive Contracts with Multiple Agents." *Journal of Economic Theory* 33 (1): 152–171.
Dimos, C., and G. Pugh. 2016. "The Effectiveness of R&D Subsidies: A Meta-Regression Analysis of the Evaluation Literature." *Research Policy* 45 (4): 797–815.
Dorie, V., J. Hill, U. Shalit, M. Scott, and D. Cervone. 2019. "Automated Versus do-it-Yourself Methods for Causal Inference: Lessons Learned from a Data Analysis Competition." *Statistical Science* 34 (1): 43–68.
Fabrizio, K. R., and U. Tsolmon. 2014. "An Empirical Examination of the Procyclicality of R&D Investment and Innovation." *Review of Economics and Statistics* 96 (4): 662–675.
Franklin, J. M., J. A. Rassen, D. Ackermann, D. B. Bartels, and S. Schneeweiss. 2014. "Metrics for Covariate Balance in Cohort Studies of Causal Effects." *Statistics in Medicine* 33 (10): 1685–1699.
García-Quevedo, J. 2004. "Do Public Subsidies Complement Business R&D? A Meta-Analysis of the Econometric Evidence." *Kyklos* 57 (1): 87–102.
Geroski, P. A. 1995. "Do Spillovers Undermine the Incentive to Innovate?" In *Economic Approaches to Innovation*, edited by S. Dowrick, 76–97. Aldershot: Edward Elgar.
González, X., and C. Pazó. 2008. "Do Public Subsidies Stimulate Private R&D Spending?." *Research Policy* 37 (3): 371–389.
Hainmueller, J. 2012. "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis* 20 (1): 25–46.
Hainmueller, J., and Y. Xu. 2013. "Ebalance: A Stata Package for Entropy Balancing." *Journal of Statistical Software* 54 (7): 1–18.
Hall, B. 1992. "Investment and Research and Development at the Firm Level: Does the Source of Financing Matter?" National Bureau of Economic Research Working Paper No. 4096.
Hall, B. 2002. "The Financing of Research and Development." *Oxford Review of Economic Policy* 18 (1): 35–51.
Hall, B. 2010. "The Financing of Innovative Firms." *Review of Economics and Institutions* 1 (1): 1–30.
Hassine, H., and C. Mathieu. 2020. "R&D Crowding out or R&D Leverage Effects: An Evaluation of the French Cluster-Oriented Technology Policy." *Technological Forecasting and Social Change* 155 (June): 1–13.
Heckman, J. J., H. Ichimura, and P. Todd. 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65 (2): 261–294.
Hirano, K., G. W. Imbens, and G. Ridder. 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica* 71 (4): 1161–1189.
Hirshleifer, J. 1961. "Risk, the Discount Rate, and Investment Decisions." *American Economic Review* 51 (2): 112–120.
Hud, M., and K. Hussinger. 2015. "The Impact of R&D Subsidies During the Crisis." *Research Policy* 44 (10): 1844–1855.
Imai, K., G. King, and E. Stuart. 2008. "Misunderstandings Between Experimentalists and Observationalists About Causal Inference." *Journal of the Royal Statistical Society* 171 (2): 481–502.
Kullback, S. 1959. *Information Theory and Statistics*. New York: Wiley.

Lach, S. 2002. “Do R&D Subsidies Stimulate or Displace Private R&D? Evidence from Israel.” *Journal of Industrial Economics* 50 (4): 369–390.

Laffont, J., and D. Martimort. 2002. *The Theory of Incentives: The Principal-Agent Model*. New Jersey: Princeton University Press.

Lee, C.-Y. 2011. “The Differential Effects of Public R&D Support on Firm R&D: Theory and Evidence from Multi-Country Data.” *Technovation* 31 (5-6): 256–269.

Lokshin, B., R. Belderbos, and M. Carree. 2008. “The Productivity Effects of Internal and External R&D: Evidence from a Dynamic Panel Data Model.” *Oxford Bulletin of Economics and Statistics* 70 (3): 399–413.

Mantovani, A. 2006. “Complementarity Between Product and Process Innovation in a Monopoly Setting.” *Economics of Innovation and New Technology* 15 (3): 219–234.

Marcus, J. 2013. “The Effect of Unemployment on the Mental Health of Spouses–Evidence from Plant Closures in Germany.” *Journal of Health Economics* 32 (3): 546–558.

Mcdonald, R. L., and D. R. Siegel. 1985. “Investment and the Valuation of Firms when there is an Option to Shut Down.” *International Economic Review* 26 (2): 331–349.

McMullin, J., and B. Schonberger. 2020. “Entropy-balanced Accruals.” *Review of Accounting Studies* 25 (1): 84–119.

Minton, B., and C. Schrand. 1999. “The Impact of Cash Flow Volatility on Discretionary Investment and the Cost of Debt and Equity Financing.” *Journal of Financial Economics* 54 (3): 423–460.

Mohnen, P., M. Polder, and G. van Leeuwen. 2018. “ICT, R&D and Organizational Innovation: Exploring Complementarities in Investment and Production” National Bureau of Economic Research (NBER) Working Papers, no. 25044.

Neuenkirch, M., and F. Neumeier. 2016. “The Impact of US Sanctions on Poverty.” *Journal of Development Economics* 121 (July): 110–119.

Nilsen, Ø, A. Raknerud, and D. Iancu. 2018. “Public R&D Support and Firms’ Performance: A Panel Data Study.” NHH Dept. of Economics Discussion Paper No. 13/2018.

Oster, E. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business and Economic Statistics* 37 (2): 187–204.

Pavitt, K. 1984. “Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory.” *Research Policy* 13 (6): 343–373.

Romer, P. 1990. “Endogenous Technological Change.” *Journal of Political Economy* 98 (5): 71–102.

Rosenbaum, P., and D. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrica* 70 (1): 41–55.

Salanie, B. 2005. *The Economics of Contracts*. 2nd ed. Cambridge, MA: MIT Press.

Smith, J., and P. Todd. 2005. “Does Matching Overcome LaLonde’s Critique of Nonexperimental Estimators?” *Journal of Econometrics* 125 (1-2): 305–353.

Strulik, H. 2007. “Too Much of a Good Thing? The Quantitative Economics of R&D-Driven Growth Revisited.” *Scandinavian Journal of Economics* 109 (2): 369–386.

Szücs, F. 2020. “Do Research Subsidies Crowd out Private R&D of Large Firms? Evidence from European Framework Programmes.” *Research Policy* 49 (3): 1–13.

Ugur, M., E. Trushin, and E. Solomon. 2016a. “Inverted-U Relationship Between R&D Intensity and Survival: Evidence on Scale and Complementarity Effects in UK Data.” *Research Policy* 45 (7): 1474–1492.

Ugur, M., E. Trushin, and E. Solomon. 2016b. “A Firm-Level Dataset for Analyzing Entry, Exit, Employment and R&D Expenditures in the UK: 1997–2012.” *Data in Brief* 8 (September): 153–157.

Wanzenböck, I., T. Scherngell, and M. Fischer. 2013. “How do Firm Characteristics Affect Behavioural Additionalities of Public R&D Subsidies? Evidence for the Austrian Transport Sector.” *Technovation* 33 (2-3): 66–77.

Wooldridge, J. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

Zhao, Q., and D. Percival. 2017. “Entropy Balancing is Doubly Robust.” *Journal of Causal Inference* 5 (1): 1–19.

Zúñiga-Vicente, Á, C. Alonso-Borrego, F. Forcadell, and J. Galán. 2014. “Assessing the Effect of Public Subsidies on Firm R&D Investment: A Survey.” *Journal of Economi•c Surveys* 28 (1): 36–67.