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

Mehmet Ugur[a] and Eshref Trushin[b]

[a] *University of Greenwich Business School;* [b] *de Montfort University, Department of Economics*

Abstract

Drawing on the theory of contracts and Schumpeterian models of innovation, we argue that direct public support for business R&D may deliver sub-optimal outcomes if firms are risk-averse and have private information about their R&D productivity. Using observable proxies for risk aversion and R&D productivity, we report that the average treatment effect (ATT) in the sample of sample of 43,650 British firms is positive but highly heterogenous. The ATTs tend to be: (a) *insignificant or negligible* when the perceived risk of R&D investment is high due to crisis episodes or because of investment in basic research; (b) *insignificant* among larger and older firms and firms closer to the R&D frontier; and (c) *positive and larger than the average* among small and young firms and firms further away from the R&D frontier. Our findings point out to conundrums in the use of R&D subsidies as an innovation policy tool: The case for R&D subsidies is stronger during economic downturns, when the investment is in basic R&D and when firms have a higher probability of innovation success; but the subsidy is less likely to increase business R&D under these conditions.

Key words: Treatment effect, R&D subsidy, innovation, additionality, entropy balancing, contract theory, Schumpeterian models.

* Paper presented at the Royal Economic Society Conference, 12 April 2021.
Based on research funded by the ESRC under grant ES/K004824/1. 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. For information about the datasets, see Office for National Statistics (2019a; 2019b).

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

1. Introduction

The case for public funding of business R&D is based on the existence of an *R&D gap*, which reflects the extent to which a firm's actual level of R&D investment is below the socially optimal level because of market failures. One type of market failure is due to imperfect appropriability of the innovation benefits (Arrow, 1962; 1996; Romer, 1990). The other results from financial market imperfections. The latter are exacerbated by the severity of the information asymmetry, adverse selection and moral hazard problems; and impose a financial constraint on innovative firms (Bloom et al., 2007; Czarnitzki and Toole, 2007; Hall, 1992; 2002; 2010; Minton and Schrand, 1999). The insights from these theoretical perspectives enjoy considerable empirical support and have long informed policy choices in favour of public subsidies for business R&D investment.

Nevertheless, economic theory also suggests that the R&D gap depends on the balance between two opposite effects: the investment-deterring effect of knowledge spillovers and capital-market imperfections *versus* the investment-inducing effects of creative destruction and market-stealing dynamics unleashed by intra-industry innovation (Bloom et al., 2019).. Secondly, knowledge spillovers and R&D investment may be complementary or substitutes. Complementarity is more likely if firms must invest in R&D (build up absorptive capacity) to benefit from knowledge externalities (Cohen and Levinthal, 1989; Geroski, 1995a; Branstetter and Sakakibara, 1998). Hence, the effects of R&D subsidies on business R&D investment would differ, depending on the balance the deterrence and inducement effects of investment in innovation and the extent to which absorptive capacity is necessary for benefiting from R&D externalities. More importantly, however, public support would have heterogeneous effects on firm effort (including R&D investment), depending on the level of information asymmetries between the funder and funded firms and the risk aversion of the latter (Laffont and Martimort, 2002; Akcigit et al., 2019). Risk aversion is also an important determinant of R&D investment in Schumpeterian and third-generation models of innovation and growth (Aghion et al., 2014; 2015; Akcigit and Kerr, 2018; Strulik, 2007), where R&D intensity decreases with the firm's risk aversion at each level of R&D profitability.

Effect-size heterogeneity has been acknowledged and discussed in the policy evaluation literature reviewed below. However, the existing explanations have remained fragmented; and reflect a tendency of over reliance on the incentive effects of R&D subsidies through cost reduction or relaxation of the financing constraint. Furthermore, there is a tendency to overlook the implications of information asymmetries between the funders and funded firms; and whether the responsiveness of the R&D investment to policy interventions varies by different R&D types associated with different levels of return uncertainty. The aim of this paper is to accord theory a more prominent role in predicting and explaining the variations in the effect of R&D subsidies on business R&D investment.

To achieve this aim, we first draw on theory of contracts (Laffont and Martimort, 2002; Akcigit et al., 2019) to demonstrate why the effect of subsidy on business R&D investment is a second-best outcome, why subsidy allocations may remain sub-optimal, and how information asymmetry and risk aversion impinge on the relationship between R&D subsidies and business R&D investment. Then, we draw on Schumpeterian and third-generation models of innovation and growth (Aghion et al., 2014; 2015; Akcigit et al., 2019; Strulik, 2007) to demonstrate how theoretical constructs such as R&D gap, R&D productivity and risk aversion that are unobservable for the funder and the researcher can be proxied by observed variables related to firm age, size and R&D intensity, R&D type and downturns in the business cycle. Finally, we utilise a novel treatment-effect estimator based on the entropy balancing (EB), which ensures a high degree of covariate balance between the treatment and the control groups up to three moment conditions: mean, variance, and skewness (Hainmueller, 2012).[1]

Our findings provide strong and consistent support to the theoretical insights from the theory of contracts and Schumpeterian models of innovation. The ATT is *smaller* when R&D investment is associated with higher return volatility/uncertainty, as it is the case when the investment is in basic R&D or it is undertaken during periods of the *dot-com* crisis from 2000 to 2002 or the global financial crisis from 2008 to 2010. The ATT for all types of R&D inputs is positive in the full sample, but this conceals a high degree of heterogeneity. The effect is

[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. 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).

positive and larger than the average among young and small firms and firms further away from the R&D frontier, which receive only 10 percent of the total subsidy or less; but it is *insignificant* among older and larger firms and firms closer to the R&D frontier, which receive 90 percent of the total subsidy or more. Effect-size heterogeneity with respect to firm type holds for all R&D input types considered.

The findings and their robustness to a range of sensitivity checks reveal inherent conundrums in public support for business R&D. On the one hand, it appears socially optimal to provide public support for basic research, during economic downturns, and when firms are more likely to convert R&D inputs into profitable innovations. On the other hand, public subsidies are less likely to generate additionality effects under these conditions because: (i) the firms' R&D investment is less sensitive to the subsidy when *R&D return uncertainty* is high either due to low market readiness of basic R&D or because of increased product market uncertainties associated with crisis periods; and (ii) large and old firms and firms closer to the R&D frontier have *narrower R&D gaps* and are more likely to extract informational rents by concealing their true types in terms of R&D productivity and R&D gap.

The rest of the paper is organised as follows. In section 2, we review the relevant literature to document the extent of heterogeneity in reported effect-size estimates and make the case for an *ex ante* theoretical framework that can shed light on the sources of heterogeneity in the empirical findings. Section 3 discusses the funding regime in the UK, spells out the information asymmetries it may entail, and draws on the theory of contracts and Schumpeterian models of innovation to demonstrate why R&D return uncertainty and firm age, size and distance to R&D frontier can be used as predictors of the variation in the subsidy's effects on R&D investment. In section 4, we present our dataset and empirical strategy. We first provide evidence on the treated and untreated samples, the percentage of firms in receipt of government subsidy, the funding rate, and the distribution of the public subsidy by firm age and size deciles. Then we discuss the EB methodology and the wide range of pre-treatment covariates that we balance to eliminate bias in the estimations of the treatment effects from observational data. The empirical results are presented in section 5, complemented with additional sensitivity and balancing checks in the on-line Appendix.

2. Effect-size heterogeneity in the literature

Following the pioneering work by Blank and Stigler (1957), a growing number of studies have utilised a variety of datasets and estimation methods to establish whether public support for business R&D has complementary or substitution effects. David et al. (2000) review 14 firm-level studies, of which three reported additionality effects, five studies based on US data reported crowding-out effects, and the remaining six reported mixed findings. Additionality effects are more likely to be reported by studies based on continental European data. This is the case in Czatrnitzki and Lopes-Bento (2013) on the effect of R&D subsidies in Flanders; Czarnitzki and Hussinger (2004) and Hussinger (2008) in Germany; Duguet (2004) in France; and Aerts and Schmidt (2008) covering both German and Flemish firms. However, effect-size estimates are heterogeneous even among studies using European data (Czarnitzki and Toole, 2007; Hud and Hussinger, 2015; Takalo et al., 2013).

Effect-size heterogeneity is related to firm size and age in several studies. For example, Marino et al., (2016) investigate the effect of subsidies on private R&D expenditures in a sample of French firms and report that the effect varies by the subsidy rate and by firm size. The crowding-out effect is more prevalent among small firms and firms in receipt of large and medium-sized subsidies. In contrast, Gonzalez and Pazo (2008) and Lach (2002) report that small firms are more likely to be associated with additionality effects. Larger additionality effects among small firms is also reported in a more recent study by Venino et al., (2019), who find that R&D subsidies have larger output (employment and turnover) additionality effects among smaller firms. A similar finding is reported by Nilsen et al., (2018), who find that the output (value-added and revenue) additionality effects are larger among start-up firms compared to older firms. Crowding-out or weaker additionality effects among older firms are also reported in Aristei et al., (2017). The difference in the treatment effect is usually explained by differences in the severity of the financing constraint, which is assumed to be more biting among small or young firms (Czarnitzki and Hottenrott, 2011; Zúñiga-Vicente et al., 2014).

R&D type is another source of effect-size heterogeneity reported in the literature. For example, Aerts and Thorwarth (2008) report that R&D subsidies tend to reduce firm investment in basic as opposed to development research. In contrast, Clausen (2009) and Czarnitzki et al., (2011) report that R&D subsidies tend to have stronger additionality effects on investment in basic research; and the effect on development research expenditures is either negative in the former or insignificant in the latter study. It seems that R&D subsidies are ineffective in generating additionality in the case of development R&D investment because firms already have stronger incentives to undertake such investment in projects that are closer to the market, face less return uncertainty, and have higher tangibility.

Given this background, effect-size heterogeneity has been acknowledged by both narrative reviews and meta-analysis studies. For example, Garcia-Quevedo (2004) reviews 28 studies that utilize firm-level data and reports that seven studies find additionality, 10 studies find no significant effects, and 11 studies report crowding-out effects. In a meta-analysis study, Dimos and Pugh (2016) synthesize 660 effect-size estimates from 52 primary studies and report that the additionality effect that can be identified is positive but small and conceals a high degree of heterogeneity. Crowding-out or insignificant effects constitute about 45% of the evidence base, with the remaining 55% indicating additionality effects. A similar picture emerges from a systematic review of 168 effect-size estimates from 77 studies by Zúñiga-Vicente et al. (2014), where 40% of the reviewed findings indicate crowding-out or insignificant effects and 60% indicate additionality effects.

Zúñiga-Vicente et al., (2014) observe that the studies tend to focus on method development and exploring different country samples or time periods. As a result, the explanations for observed heterogeneity have remained patchy and unsystematic. To address this gap, the authors offer a number of assumptions about why the effect of R&D subsidies should be expected to vary by observable sources of heterogeneity such as the firm's subsidy history, the time lag, the existence of financial constraints, the composition of the R&D investment, and the generosity and sources of the public subsidies. Their overall conclusion is that future research should be more systematic in modelling, estimating, and explaining these sources of heterogeneity to develop a better understanding of where and when R&D subsidies may or may not induce additionality effects.

So far, four studies have attempted to identify the sources of the heterogeneity in the treatment effects in a systematic manner. Czatrnitzki and Lopes-Bento (2013) is an empirical attempt at investigating whether the treatment effect varies over time, or by the number of grants received, or by multiplicity of the funding sources. The authors regress the treatment effects obtained from a propensity score estimator on dummy variables that capture the potential sources of heterogeneity, and report that the effects do not vary over time or with the number of grants received or funding sources. Lach (2002) hypothesize that the subsidy's effect on business R&D investment depends on innovation success probability and the marginal cost of hiring R&D personnel. If the funder prefers to fund projects with higher success probability, the scope for additionality effects is reduced because such projects might have been undertaken even without a subsidy. Furthermore, subsidised projects may crowd-out the financing of unsubsidised projects within the same firm if the marginal cost of hiring additional R&D personnel is high. Drawing on firms in Israel, Lach (2002) reports that the funding regime fails to create additionality because most subsidies are awarded to projects submitted by large firms, which would undertake those projects without the subsidy.

Lee (2011) draws attention to the importance of firm size and argues that R&D subsidies are less likely to generate additionality effects among large firms. Crowding-out among large firms is explained by their proximity to the technology frontier, which leaves little need for catching up. In contrast, additionality effects among small firms are explained by their distance to the frontier and the need for catching up. Finally, Wanzenböck et al., (2013) identify three firm characteristics as potential sources of heterogeneity in behavioural additionality: R&D intensity/experience, technological specialization, and collaboration propensity. Their findings indicate that R&D-intensive firms are less likely to exhibit behavioural additionality because such firms have the capacity to undertake the desired R&D projects irrespective of public support.

An issue common to these studies is that the theoretical framework that informs their testable hypotheses is either partial (i.e., it covers only a subset of the factors that moderate the relationship between the subsidy and the firm's R&D effort) or it is developed in an eclectic manner - for example by juxtaposing the technological characteristics and collaboration propensities of the firms. We aim to transcend these limitations by focusing on the subsidy contract as a principal-agent setting and trace the implications of asymmetric information and risk aversion in that setting for the for the firm's response to the R&D subsidy. The proposed approach is versatile enough to generate theoretically informed hypotheses on how the effect of the subsidy on business R&D investment is moderated by a wide range of firm characteristics, R&D types and the timing of the policy intervention. The moderating factors we identify include but transcend the factors proposed in the existing work.

3. From the subsidy contract to R&D effect heterogeneity: The role of risk aversion and information asymmetry

Our sample of firms receive direct public support form UK government departments, their agencies, and non-departmental public bodies like the *Technology Strategy Board*, including its successor, *Innovate UK*.[2] They also receive funding from the European Union (EU) commission.[3] Despite the involvement of multiple funders, two main features of the UK subsidy regime stand out. First, the largest part of the subsidies has been managed by non-departmental agencies, of which *Innovate UK* is the incumbent. Secondly, the UK support for business R&D has to comply with the EU's state-aid rules, under which R&D grants should not lead to unfair competition. The risk of unfair competition is measured by the proximity of the applicant's project to its market operations – the so-called market readiness level (MRL). R&D activities that score 1 on the MRL scale such as basic research are furthest away from the market and qualify for public funding of up to 100% of the project costs (Table 1). The funding rate gradually declines for R&D activities closer to the market and varies between 25%-70% of the project cost, depending on firm size.

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].

To secure public funding, the applicant must demonstrate: (i) whether the project could be undertaken without public funding; (ii) the extent to which the project represents value for money for the taxpayer; and (iii) how the applicant will benefit from the innovation, including the latter's impact on productivity and growth. In addition, the applicant must provide

[2] 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.

[3] We investigate only the effects of UK subsidies on business R&D investment. To ensure identification of the UK subsidy's effect, we obtain balance between treated and control firms with respect to whether the firm receives EU funding in addition to a wide set of covariates.

information about the scope for additionality by explaining why private finance may not be available for the project.[4]

The firm's planned R&D investment is observable to the funder because funding is conditional on the firm's supply of information investment plans and project implementation within those plans. However, information asymmetry *does* exist with respect to three criteria important in the theory of contracts. One is the *risk aversion* in the firm's decision-making for R&D investment, which determines the discount rate that the firm utilizes to select between R&D projects. The second is the firm's *R&D productivity* defined as the success with which the firm converts R&D inputs into innovation outputs. The third is the firm's *R&D gap* defined as the difference between the firm's observed/actual level of R&D investment and the socially optimal level. Furthermore, the funder takes the existing intellectual property regime as given and does not monitor the firm's price-costs margin.

In the theory of contracts (Salanie, 2005; Laffont and Martimort, 2002), the consequences of the principal-agent relationships are analysed under the following behavioural assumptions: (i) the principal and the agent are rational and both maximize their utilities; (ii) the agent has private information about its type in terms of own productivity/efficiency, but the funder knows only the distribution of the agent types; and (iii) the principal moves first by offering R&D subsidy, while the agent accepts it if the offer satisfies its participation constraint.

Under perfect information about the firm's R&D productivity and perfect competition in the product market, the optimal (the first-best) subsidy that would maximise social welfare is equal to a 'Pigouvian correction' for positive externalities (spillovers) from the firm's own innovation. In this first-best scenario, the subsidy rate is just enough to correct for the incentive problem caused by imperfect appropriability of the innovation benefits. Stated differently, the funder does not have to offer an additional 'screening premium' to tease out information about the firm's R&D productivity (hence its R&D gap) as this information is common knowledge. Additionally, the funder does not have to provide a 'monopoly quality valuation correction' premium because the firm is operating at the socially optimal level of output in a competitive industry.

[4] See Innovate UK, General guidance for grant applicants at https://www.gov.uk/guidance/innovate-uk-funding-general-guidance-for-applicants.

Under asymmetric information, however, the funder must augment the ‘Pigouvian correction’ with a ‘screening term’ aimed at inducing the high-productivity firms to reveal their true types (Akcigit et al., 2019). This Nash equilibrium outcome is determined by the participation constraint of the efficient firm and indicates a deviation from the first-best level of the subsidy. The funder satisfies the participation constraints of the high-R&D productivity firms to minimise the distortion of the subsidy allocations in favour of low-R&D-productivity firms. In the theory contracts, the magnitude of the screening term is an increasing function of the proportion of low-R&D-productivity firms in the applicant pool (Akcigit et al., 2019).

When agents are heterogenous and correlation between their types is unknown, the high-R&D--productivity firm can mimic the low-productivity type and extract informational rents. This is because *observed* R&D investment would be only a noisy indicator of the firms’ *unobserved* R&D effort and productivity. Both Laffont and Martimort (2002) and Akcigit et al. (2019) demonstrate that monitoring an *observed* indicator of the *unobserved* effort does not enable the principal to elicit truthful information about the agents’ true type. 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 et al., 2019). Taken together, these deviations reflect a trade-off between efficiency gains from allocating subsidies to successful innovators and potential failure to induce additional R&D effort by high-R&D-productivity firms that would invest in R&D even in the absence of the subsidy.

A third source of deviation from the first-best outcomes 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 higher level of incentives that is positively related to (i) the firm’s risk-aversion and (ii) the marginal cost of the effort needed to deliver the outcome expected by the funder (Laffont and Martimort, 2002). Therefore, it is may be socially optimal to offer R&D subsidies aimed at closing the R&D gap, but the intervention may not generate the desired outcome if risk-aversion induces the innovative firms to discount the future returns from R&D investment at higher rates.

Overall, the theory of contracts predicts that information asymmetry between the funder and the funded firms is conducive to the extraction of informational rents by high-R&D-productivity firms; the deviations from first-best outcomes would be exacerbated as the firms' risk aversion increases. A pertinent question here is whether the funder can achieve first-best outcomes by monitoring the firms' observable performance, which is the actual level of R&D investment. Demski and Sappington (1984) demonstrate that this is feasible only if the agents' productivities are correlated. If the agents are heterogenous, the high-productivity agent can still mimic the low-productivity type and extract informational rents. Furthermore, the *observed* R&D investment is a noisy outcome of the firms' *unobserved* R&D productivity. Both Laffont and Martimort (2002) and Akcigit et al. (2017) demonstrate that monitoring the *observed* outcomes of *unobserved* agent productivity or effort does not enable the principal to elicit truthful information about the agent's true type.

We now draw on Schumpeterian models of innovation (Aghion et al., 2014; 2015) to identify *observable* firm types and R&D types that can be mapped onto *unobserved* R&D productivity and risk aversion as factors that moderate the effect of R&D subsidy on business R&D investment. In these models, firms survive and grow as they add new product lines; or shrink and eventually exit as their product lines become less profitable or obsolete due to creative destruction. Stated differently, firm value and age are related positively to the firm's R&D productivity, defined as the success with which the firm converts R&D investments into innovative and profitable product lines. Following Aghion et al., (2014), we can state the normalised average value of the innovative product as follows:

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

The value of the innovative product line in year t, v_t, is increasing with adjusted profitability of the innovative product line (π_{At}), which is the difference between gross profits (π_t) and the cost of R&D investment (C_t^{RD}). It 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. Furthermore, the firm's market value is a linear function of the number of innovative product lines (k_t), the latter's normalised average value (v_t) and output per product line (Y_t), as indicated in (2) below.

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

In equation (3), the firm chooses the R&D intensity (z_{it}) that maximises the contribution of the innovative product line to firm's market value. Here, $Wage\_cost_t$ is the cost of employing R&D scientists and technicians, and $1/\eta$ is the elasticity of innovation with respect to employment of scientists and technicians.

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

R&D intensity in (3) is increasing with R&D productivity (profitability of the innovative product line); but it is decreasing with the discount rate (ρ_t), the rate of creative destruction and wage cost. Finally, the firm's survival time is a positive function of its R&D productivity, as demonstrated in Ugur et al. (2016a). Assuming that the firm's market value follows a Wiener process until liquidation (McDonald and Siegel, 1985), Ugur et al. (2016a) relates survival time to R&D productivity as indicated in (4) below, where $E[t]$ is expected time until exit and V_0 is the initial market value of the firm.

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

Three predictions follow from equations 1 – 4 above. First, firms that are successful in converting R&D investment into innovative product lines have larger market values (equation 2) and survive longer (equation 4). Secondly, the R&D intensity is higher when the firm is more successful in converting R&D inputs into innovative and profitable product lines (equation 3). Therefore, high-R&D-productivity firms are closer to the R&D frontier and as such would have narrower R&D gaps – i.e., the difference between their actual and socially-optimal levels of R&D investment would be smaller.

Finally, more risk-averse firms would utilise a higher discount rate, ρ_t, and invest less in R&D at each level of R&D productivity (equation 3). This is in line with neo-classical and evolutionary models (Arrow, 1962; Nelson, 1959), where return uncertainty is conducive to lower demands for R&D investment under risk aversion. It is also in line with findings from the real option theory of investment, where return uncertainty increases the value of the waiting option for costly and irreversible investments (Abel and Eberly, 1996; Abel et al., 1996).

Other findings that support the prediction from the Schumpeterian models include Kwon (2010), who investigates how firms allocate resources when they compete for multiple patents with heterogeneous research projects simultaneously. This work demonstrates that the firms' resource allocation is biased away from risky and basic research, even when spillovers do not exist; and the market may fall short of supplying major innovations despite large R&D expenditures and strong patent protection. Similar support can be seen in empirical work, which report that higher R&D return uncertainty is associated with lower R&D investment when the firms are risk averse. (Goel and Ram, 2001; Caggese, 2012; Ghiglino and Tabasso, 2016).

In what follows, we consider the three predictions above in reverse order and develop three hypotheses that we aim to test in this study. Starting with the role of risk aversion, we hypothesize that an increase in R&D return uncertainty induces the firm to utilise a higher discount rate and leads to lower business R&D investment that, in turn, leads to a wider R&D gap. To induce the firm to close the R&D gap, the funder must offer a higher subsidy rate that includes a risk premium. This is indeed what we observe in the UK funding criteria that fund basic research up to 100% of the project cost (Table 1) and in the subsidy allocations that cover a higher proportion of the eligible firms during crisis periods (Table 2).

Although higher subsidy rates are required to satisfy the participation constraint of the risk-averse firms, the latter's R&D investment are less responsive to subsidy when R&D returns are uncertain (Aristei et al., 2017; Czarnitzki and Toole, 2013; Bloom, 2007). This is also what is implied by equation (3) above: the subsidy reduces the cost of R&D, increases the adjusted profits, and induces the firm to increase its privately optimal R&D intensity; but the subsidy's effect on the firm's R&D intensity would be small or insignificant if the firm discounts the future R&D returns at higher rates.

This is in line with empirical findings indicating that the responses of both subsidised and unsubsidised firms to financial crises are pro-cyclical (Fabrizio and Tsolmon, 2014). It is also in line with findings that both subsidised and unsubsidised firms invest less in basic R&D due to higher return uncertainty (Nelson, 1959; Czarnitzki and Toole, 2013). Finally, the insight from the Schumpeterian models is also consistent with theoretical findings on increasing returns to 'waiting'. 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 et al., 2007; Bouvatier and Lepetit, 2008). Hence, we state our first hypothesis (**H1**) as follows:

H1: *Compared to non-crisis periods or to R&D types with higher market readiness levels, R&D subsidies are less effective in generating additionality effects during financial crises and when the investment is in basic R&D.*

If supported by evidence, **H1** implies that the financial constraints faced by innovative firms (Czarnitzki and Toole, 2007; Hall, 2002; 2010; Minton and Schrand, 1999) may justify the granting of R&D subsidies but the relaxation of the financial constraint through subsidies does not necessarily induce business R&D investment. This is because the R&D subsidy addresses market failures due to risk aversion, moral hazard and information asymmetry between the firms and capital markets; but creates new problems that arise from information asymmetry and risk aversion in the relationship between the firms and the public funder.

Insights from the theory of contracts and Schumpeterian models are better placed to explain not only why the case for R&D subsidies is stronger when capital/financial market failures are severe and, hence, the R&D gaps are wider. They can also explain why subsidies may remain ineffective in bridging the R&D gap when risk-averse firms face higher return uncertainty. In addition, insights from the theory of contracts imply that the funder may have little power to nudge the firms towards the socially optimal level of R&D investment because the subsidy contract does not provide for a third-party adjudicator that can "…credibly impose punishments" on the party that violates the contract (Laffont and Mortimort, 2009, pp.253,271).

Our second hypothesis (**H2**) relate to the moderating role of the firm's proximity to the R&D frontier. As indicated above, Schumpeterian models of innovation predict that, *ceteris paribus*, an increase in R&D productivity rises the average value of the innovative product line in (1) and this leads to a higher level of R&D intensity, z_i, in (3). On the other hand, the theory of contracts predicts that the funder must satisfy the participation constraint of the high-R&D-productivity firms to minimise the distortion of the subsidy allocations in favour low-R&D-productivity firms. However, information asymmetry enables the high-R&D-productivity firms to conceal their types and extract informational rents (Laffont and Mortimort, 2009; Akcigit et al., 2017_. Given these insights, we state our second hypothesis (**H2**) as follows:

H2: *The responsiveness of different types of R&D investment to public subsidies is weaker the closer is the firm's R&D intensity to the R&D frontier in the industry.*

H2 is consistent with and encompasses diverse findings in the empirical literature on R&D subsidies. For example, Lach (2002) observes that funders prefer to fund projects with higher success probability, but such preference reduces the scope for additionality effects because such projects typically have higher private rates of return and might have been undertaken without a subsidy. Similarly, Wanzenböck et al., (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 desired R&D projects irrespective of public support. Finally, Lee (2011) reports that crowding-out effects among large firms are due to their proximity to the technology frontier, which leaves little need for catching up.

Our third hypothesis (**H3**) is based on mapping the unobserved R&D productivity and R&D gap on to firm age, size, and market share. In equation (1), the higher is the R&D productivity, the larger is the value of the innovative product line (v) that, in turn, determines firm size through aggregation of product lines. Firms with higher R&D productivity have a larger number of profitable product lines (larger size) and larger market values. These firms, in turn, would survive longer in accordance with (4). Therefore, 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 be less responsive to R&D subsidy compared to other firms with opposite characteristics. Hence, 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 across all R&D types.*

The analysis above and the derived hypotheses offer a systematic approach to explaining why effect-size estimates vary between and within studies (Zúñiga-Vicente et al., 2014). 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.

4. Data and empirical strategy

4.1 Data

Our data is from the Business Research and Development Database (BERD) and Business Structure Database (BSD).[5] The BERD survey is based on a sample of R&D-active firms stratified by product group and employment size-bands. The stratified sample consists of large firms (size-band1) with 400+ employees (sampled 1:1); size-band2-firms (100-399 employees) sampled 1:5 and size-band3-firms (1-99 employees) sampled at a rate of 1:20. In 2012, 400 large R&D-spenders, that are included in the survey every year, account for 78% of UK business R&D expenditures (ONS, 2012, p.14). The survey questionnaire asks reporting units to state intramural (in-house) and extramural (contracted-out) R&D expenditures. For intramural R&D, the firm is also asked to provide a breakdown by current and capital R&D expenditure. In turn, 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 …")[6]. Finally, the questionnaire asks the firm to state the amount of the intramural R&D financed privately, from UK public funds and from EU funding.[7]

In the dataset, the firm's subsidy (treatment) status in each year is determined by whether the firm receives UK funding in that year. Moreover, the UK funding disbursed in any year is conditional on R&D expenditures *incurred* in the implementation of funded R&D project(s) during the year. Therefore, the level of R&D investment by the firm in any year includes both the subsidy received and the expenditures financed privately. As such, the dataset allows for identifying the treatment's (subsidy's) effect on various components of the firm's R&D investment in each year - provided that the treated and untreated samples are comparable with respect to all relevant characteristics apart from the treatment status.

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.

[5] Office for National Statistics (2019a; 2019b). See also Ugur et al. (2016b).
[6] See examples of the questionnaires at https://www.ons.gov.uk/businessindustryandtrade/business/
[7] 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.

The BSD contains information on firm turnover, employment, age, survival status, etc. Combining information from both datasets, we obtain a sample of 43,650 firms observed from 1998 to 2012. Table 2 presents annual information about R&D intensity (R&D a percentage of turnover), subsidy rates (UK subsidies as a percentage of privately-funded R&D), and coverage ratios (percentage of firms subsidised) for each year. Column 1 indicates that privately funded R&D intensity has fallen in the final year of the *dot.com* crisis in 2002 and from 2009 onwards during the global financial crisis. This is in line with patterns reported in the relevant theoretical and empirical literature (Acemoglu and Linn, 2005; Dubois *et al*., 2015; and Aghion *et al*., 2012). In contrast, the subsidy rate (column 2) and the coverage rate (column 3) have increased from the second year of the crisis periods and remained higher than average for at least two years thereafter.

Table 2: Private R&D expenditures and UK subsidies by year.

Year	**1. Private R&D intensity** (Private R&D expenditures as % of turnover)	**2. Subsidy intensity** (Subsidy as % of private R&D expenditures)	**3. Coverage** (Subsidised firm/years as % of total firm/years)
1998	2.7	10	86
1999	2.9	10	86
2000	2.8	9	77
2001	2.9	7	94
2002	2.5	4	94
2003	3.4	12	97
2004	2.9	11	95
2005	2.9	8	90
2006	2.4	7	92
2007	2.3	6	89
2008	2.8	6	84
2009	2.2	7	97
2010	2.2	7	95
2011	2.5	8	96
2012	2.3	7	97
Average	**2.6**	**8**	**92**

Note: Pooled sample of 43,650 firms with 154,980 firm/year observations over 1998-2012. Excludes firm/year observations with private R&D intensity greater than 1.

The evidence in Table 2 suggests that the funder has increased the level of support after the crisis events, perhaps with a view to encourage R&D investment when the firm's perceived risks are higher and the financing constraint is more biting due to the downturn in the business cycle. These funding decisions may be justified from a social-welfare perspective but, under H1, we expect their additionality effects to be smaller than the effects in the full sample.

Table 3: Private R&D expenditures and UK subsidies by age and size deciles.
(Pooled panel of 43,650 firms with 154,980 firm/year observations)

Panel A - By age deciles	Private R&D (£ bn.)	Private R&D intensity (Private R&D as % of turnover)	Subsidy (£ bn.)	Subsidy rate (Subsidy as % of private R&D)	Coverage (Subsidized firm-years as % of total 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%**		
Share of top 30%	**59.8%**		**63.4%**		
Share of top 10%	**30.47%**		**23.37%**		
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%**		
Share of top 30%	**96.7%**		**96.9%**		
Share of top 10%	**88.23%**		**93.32%**		

Note: Excludes firm/year observations with private R&D intensity greater than 1.

Table 3 provides further descriptive information by age and size (employment) deciles. The percentage of the subsidy allocated to firms in the top 50% of the distribution is proportionately higher, at 85.3% and 98.4%, respectively. The skew in favour of old and large firms is even more evident in the top decile of the distribution, where 30% of the subsidy is allocated to oldest and 88% is allocated to largest firms. 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.

Key characteristics of the subsidized and non-subsidized samples are summarised in Table 4.[8]

Table 4. R&D intensity and firm characteristics by treatment status.

	Non-subsidized (Untreated)		Subsidized (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 to 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.

Compared to non-subsidised counterparts, subsidized firms spend less on R&D (row 2) and employ less R&D personnel (row 6). The subsidized firms are also smaller than the non-subsidized firms in terms of turnover (row 9) and total employment (row 10). Nevertheless, in

[8] 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 et al., 2016a).

terms of the R&D input intensity, subsidized firms have a relatively higher R&D intensity (rows 3 and 4) and higher R&D personnel intensity (row 7) compared to non-subsidized firms.

The proportion of start-ups (row 11) and young firms (row 12) in the subsidized sample is higher than their proportion in the non-subsidized sample. In contrast, the proportion of mature firms (row 13) and old firms (row 14) is higher in the non-subsidized sample. However, the difference between the proportions of old and mature firms in the subsidized and non-subsidized 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 subsidized and non-subsidized samples.

4.2 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. To achieve this aim, we start by addressing the issue of alignment between the policy intervention and the R&D investment outcomes.

In the data, the firm is treated if it receives R&D subsidy for any project type, which can be a basic, development, experimental or capital R&D project. As such, the data does not allow for pairing a specific subsidy (e.g., a subsidy for a basic R&D project) with a specific R&D type (e.g., basic R&D expenditure). This is a common data constraint problem that affects most of the literature on the effects of the subsidy as a direct support instrument and all of the literature that investigates the effectiveness R&D tax credits as a direct support instruments. The problem is addressed by assuming (often implicitly) that different types of R&D expenditures (e.g., basic, development, experimental, or capital R&D) are complements. Under this assumption, a subsidy received or R&D tax credits claimed for a particular R&D project (e.g., a basic R&D project) affect the firm's R&D expenditure on basic R&D as well as other R&D expenditures that may be necessary either to implement the basic R&D project or exploit the new knowledge from basic research. A similar complementarity is assumed when the firm receives subsidies for multiple R&D projects, the subsidisation of which affects R&D spending on those projects and the remaining projects that may be necessary for successful implementation of the subsidised projects.

Nevertheless, there is no theoretical consensus on whether R&D types are complementary or substitutes. In a seminal paper on the relationship between R&D capital and productivity, Griliches (1979) indicates that privately-funded and publicly-funded R&D capital can be either substitutes or complements; and recommends that the issue can be resolved only empirically. The findings that follow tend to indicate complementarity. For example, Mantovani (2006) demonstrates that, under monopolistic competition, process and product innovations are complementary and firms always prefer the simultaneous adoption of both innovation types. Using data on large R&D investors, Lokshin et al., (2008) find that extra-mural and intra-mural R&D have complementary effects on productivity, with the implication that firms that increase investment in intra-mural R&D are expected to invest more in extra-mural R&D. More recently, Mohnen *et al*. (2018) report that investments in ICT, R&D and organizational innovation are complementary, with investment in one innovation type increasing the probability of investing in others as joint investments lead to higher TFP growth than individual investments.

Given these findings, we assume complementarity between different R&D types, with the implication that the effect of a subsidy for any or all types of R&D projects in any year will affect the individual and aggregate measures of busines R&D investment in the year or thereafter in the same direction. Hence, we provide effect-size estimates for the effect of the treatment on both aggregate business R&D measures such as privately funded R&D or employment of R&D personnel; and individual R&D types such as basic research, capital R&D, development R&D, and extra-mural R&D etc. The former type of treatment effect estimation is the most common exercise in the literature. The latter is reported less frequently, but has been used to estimate the effects of a composite intervention on basic and development research investment separately (e.g., Aerts and Thorwarth, 2008; Czarnitzki et al., 2011).

Our effect-size estimator is the average treatment effect on the treated (ATT) - the conditional difference in mean outcomes for treated (subsidized) and untreated (non-subsidized) 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 of the R&D input and R&D personnel intensities – i.e., the first difference of the intensity measures in (i). Our preferred measure is the latter as first-differencing eliminates the firm-specific fixed effects. However, we conduct sensitivity checks with the logarithm of the R&D input and R&D personnel intensities.

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, it is necessary to construct a counterfactual by selecting control (untreated) firms that are as close as possible 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 et al., 1998; Smith and Todd, 2005).

The expected value of the sample average of the treatment effect on the treated (*SATT*) can be stated as follows:

$$E[SATT] = \mathrm{E}[\mathbf{Y}_{it}(1)|\ \mathbf{X}_{it-1}, D_{it}=1] - \mathrm{E}[\mathbf{Y}_{it}(0)|\ \mathbf{X}_{it-1}, D_{it}=1] \qquad (6)$$

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; $\mathbf{Y}_{it}$ is the outcome variable as defined above, with $\mathbf{Y}_{it}(1)$ indicating the outcome for the treated firm and $\mathbf{Y}_{it}(0)$ indicating the outcome for untreated firm. We use one-year-lagged covariates ($\mathbf{X}_{it-1}$) 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 $\mathrm{E}[\mathbf{Y}_{it}(1)|\ \mathbf{X}_{it-1}, D_{it}=1]$ can be estimated directly from the observed sample of treated firms, but the conditional expectation $\mathrm{E}[\mathbf{Y}_{it}(0)|\ \mathbf{X}_{it-1}, D_{it}=1]$ is the unobserved counterfactual. With randomised control trial (RCT) data, the counterfactual $\mathrm{E}[\mathbf{Y}_{it}(0)|\ \mathbf{X}_{it-1}, D_{it}=1]$ can be consistently estimated from a randomly selected control group. However, this option is not available for observational studies such as our work. The solution is to estimate $\mathrm{E}[\mathbf{Y}_{it}\ (0)|\ \mathbf{X}_{it-1}, D_{it}=1]$ using a control group of units that are equivalent to the treated units (firms) with respect to a wide range of pre-treatment characteristics.

A variety of pre-processing methods exist for estimating the counterfactual outcome with observational data. One approach is to ensure comparability between the treated and untreated samples through inverse probability weights. The other relies of propensity score weights that are used to obtain covariate balance between the treated and control groups. Both approaches estimate the ATT consistently if the propensity scores are estimated correctly to ensure that the control-group outcome is orthogonal to (independent of) the treatment status (Rosenbeim and Rubin, 1983). If this independence condition is satisfied, the counterfactual outcome can be estimated as follows:

$$E[Y_{it}(0)|D = 1] = \frac{\sum_{\{it|D=0\}} Y_{it} w_{it}^{p}}{\sum_{\{it|D=0\}} w_{it}^{p}} \tag{7}$$

Here, the correct propensity weights (w_{it}^{p}) are equal to $\frac{p_{it}}{1-p_{it}}$, where p_{it} is the correctly estimated propensity score. The weights will ensure the that the control group is balanced with treated group if the propensity scores are correct (Hirano et al., 2001; 2003).

However, propensity-score balancing methods face several challenges. First, the true propensity scores are unknown, and their estimates are model-dependent. Incorrect model specification leads to biased propensity-score and treatment-effect estimates. Secondly, it may be difficult to balance all pre-treatment covariates jointly – either because the selection into treatment is complex or the data is highly multidimensional. Such difficulties induce a process of cyclical and often *ad hoc* iterations that involve selecting a propensity score model and then checking some standardized difference for each covariate between treated and control groups. Imai et al., (2008) criticise such iterations as “propensity score tautology” with questionable statistical assumptions (King and Nielson, 2019). Cyclical iterations between matching or weighting, propensity score modeling, and balance checking often results in low overall balance (Hainmueller, 2012). Finally, propensity score methods ensure covariate balance only asymptotically even when the propensity score model is specified correctly. Remaining imbalances in finite samples require a different weighting scheme that allows for imposing restrictions not only on the first moment of the distributions, but also on higher moments such as variance and skewness (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 ($q_i = 1/n_0$). The base weight is the reciprocal of the number of units (n_0) in the control group. The EB weights thus obtained can be used to obtain the population average treatment effect on the treated (PATT) in accordance with (8).

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

The EB weights, w^{eb}, are chosen by minimizing the entropy distance *H(w)* in (6), where the base weights for the firms in the control group is q_i, subject to balance conditions for the sample moments of the control group.

$$\min_{w_i} H(W) = \sum w_i \log(w_i/q_i) \tag{9}$$

The sample moments of the control group of firms are reweighted with coefficients C^r so that it is equal to the sample moments, m_r, of the treated group of firms in accordance with (10).

$$\sum w_i C^r = m_r \tag{10}$$

The sample moment conditions can be one (mean), two (mean and variance) or three (mean, variance, and skewness). The weighting is subject to a normalization constraint that the sum of all non-negative weights is equal to one, as indicated in (11).

$$\sum w_i = 1 \text{ and } w_i > 0 \tag{11}$$

Notice that (6) minimizes a measure of dissimilarity between probability distributions of the control and treated group of firms. As such, it provides a well-known measure of informational gain from approximating one probability distribution to another. The minimization of the divergence between the probability distributions of the treated and control group is a general principal of statistical inference (Kullback, 1959). We use the EB method implemented in Stata by Hainmueller and Xu (2013), which allows for covariate balancing up to three moments (R=3): mean, variance and skewness.

We obtain EB weights for a total of 139 covariates, including: 19 pre-treatment covariates at the firm level; 7 covariates at the industry level; 8 indicator variables that capture the receipt of EU subsidies and the change in the UK's R&D tax regime; 2 crisis dummies for 2000-2002 and for 2008-2010; 15 year dummies; 4 Pavitt technological class dummies, and 84 industry dummies at the two-digit SIC level. Description and measurement of the covariates are in Table A1 in the on-line Appendix.

We estimate the ATT with weighted least squares (WLS), using EB weights as analytical weights. We account for sources of heterogeneity in two ways. Our preferred method is to estimate ATTs based on different samples that correspond to different R&D and firm types, and crisis periods. These include: (i) two crisis periods during 2000-2002 and 2008-2010; (ii) different R&D input types, including private RD, R&D personnel employment, basic R&D, applied R&D, experimental R&D, capital R&D investment, and extramural R&D; (iii) quartiles of the distributions for firm age, size, market-share, and proximity to R&D frontier; and (iv) firms in the manufacturing sector only and those that survive throughout the entire period. We also probe the robustness of the split-sample outcomes with regression analysis, where the treatment effects estimated form the full sample are regressed 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.

In the case of privately funded R&D, the ATT indicates additionality if it is positive and significant, or crowding-out in case of negative and significant value. This is because private R&D expenditures are fully funded by the firm from its own resources or through credit or equity market. An insignificant ATT estimate indicates no effect. In the case of other R&D inputs such as basic R&D, experimental R&D, or employment of R&D personnel, a positive ATT estimate indicates *either additionality or absence of full crowding-out*; whereas an insignificant or negative ATT indicates *crowding-out effect*. This is because the private and public sources of funding for these R&D expenditures are not separated.

Zhao and Percival (2017) demonstrate that the EB method: (i) is consistent with the doubly-robust estimator of the treatment effect; (ii) reaches the asymptotic semiparametric variance bound of the doubly-robust estimator when both the selection and outcome models in the latter are correctly specified; (iii) produces smaller bias than conventional doubly robust estimators; and (iv) estimates the treatment effect with the smallest variance when applied to data used in four major studies in the research field. Apart from these qualities, the EB method has the advantage of making full use of the information in the control sample - in contrast to most matching methods that omit non-matching control units. In addition, Amusa et al. (2019) find in numerical simulations that the EB tends to outperform propensity-score-based matching estimators. As such, EB constitutes a welcome addition to the range of treatment-effect estimators that rely on direct covariate balancing instead of matching (Athey and Imbens, 2017). Finally, the method has been applied in over 40 evaluation studies (examples of which include McMullin and Schonberger, 2020; Marcus, 2013; and Neuenkirch and Neumeier, 2016) published in 11 business and economics journals.

Against these advantages, the EB weighting method has two potential limitations (Hainmueller, 2012).[9] The first arises from data quality. If the treatment and control groups are very different in size or the selection to treatment is multidimensional, the method does not provide EB weights that satisfy all moment conditions (mean, variance and skewness equivalence). The solution is either to reduce the moment conditions or obtain more data. Secondly, the EB weights for some control units may require a high level of adjustment because there may be only few 'good' control units similar to treated units. In such cases the control units may receive large weights, which increase the variance of the treatment-effect estimate.[10]

In this study, we have encountered the first issue and addressed it by restricting the moment conditions to the mean and variance, excluding skewness. Although higher moment conditions imply better covariate balance, the EB weights obtained even with one moment condition (the mean) provide better sample balancing than propensity-score weights. This is because the EB method corrects for any residual imbalances in the sample, whereas the propensity score methods ensure covariate balancing only asymptotically (Hainmueller, 2012; Hainmueller and Xu, 2013; and Zhao and Percival, 2017). Nevertheless, we take account of this limitation by augmenting the outcome model with the covariates in the weighting model to: (i) control for the effect of any remaining imbalance; and (ii) obtain more efficient ATT estimates. Our findings indicate that the augmented model *does* yield smaller standard errors, but the magnitude of the ATT estimate is not affected by inclusion of the balancing covariates.

To address the second issue, we conduct sensitivity checks to verify if large weights for some control units lead to biased treatment-effect estimates. One check consists of using base

[9] Hainmueller (2012) points out a third limitation, which is encountered when the method does not provide a weighting solution because of inconsistency in balance constraints, which we did not encounter.
[10] However, it must be indicated that both limitations also apply to other methods, including matching based on propensity scores.

weights obtained from a coarsened exact matching (CEM) routine (Iacus et al., 2011; Blackwell et al., 2009) instead of the uniform base weights (*$1/n_0$*) that are the default in the estimator. The other consists of trimming the top 1% of the EB weights and re-estimating the treatment effect with the trimmed sample. In all samples, we have found that less than 4% of the control group units are reweighted in the trimmed sample, which suggests that trimming may not be necessary (see Huber et al., 2010). Nevertheless, we have estimated ATTs trimming the top percentile of the weights and found little or no change in the estimated parameters. This was the case with or without alternative base weights. Following Hainmueller and Xu (2013), we have also checked if the variance of weights is converging to about the same minimum for a given sample, when initial weights are changed or the top percentile of the weights are omitted. The convergence is observed in all estimation samples we evaluate in this paper.

The checks above notwithstanding, we are aware that the use of pooled panel data may pose some challenges for treatment-effect estimations. First, firms may be in receipt of subsidy for several times over the analysis period. Therefore, it may be difficult to disentangle the effect of subsidy in a particular year from the effect of subsidies in previous years. We address this issue by including a past (one-year-lagged) subsidy indicator in the covariate balancing model. This ensures that the subsidy status of the firms in the year that precedes the outcome is balanced between the control and treatment groups in each sample. We also estimated the ATT by regressing the outcomes in years t+1 and t+2 on the treatment (subsidy) indicator for year t. The results indicate that the subsidy in year t has very small or insignificant effects on the outcomes in year t+1 and t+2. This is as expected because the firm receives public funding in any year only for R&D expenditures incurred in that year.[11]

Secondly, UK firms also receive European Union (EU) grants; and they are entitled to R&D tax credits (indirect support) in addition to R&D subsidies (direct support). We address these issues by including binary variables that take the value of 1 if: (i) the firm is a recipient of subsidy from UK and EU sources; (ii) the year corresponds to the implementation of the tax-credit regime as applied to SMEs in 2001; (iii) the year corresponds to the change in its implementation to include large firms in 2003; (iv) separate dummies for the R&D tax credit policy adjustments for SMEs and large firms in 2008. Inclusion of these covariates ensure that,

[11] See the discussion on the funding regime above.

in each sample, the treated and control units are balanced in terms of their exposure to the R&D tax credit policy.

Finally, we have also addressed the issue of potential time-series dependence in pooled panel data, which is usually ignored in most empirical work in the field. Overlooking the potential for time-series dependence may lead to misleading inference due to incorrect standard errors. We address this issue through bootstrapping, which resamples all time periods for each firm in the pooled data as recommended by Wooldridge (2010).

5. 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; 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 is estimated with 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 the payment of the subsidy and the implementation of the subsidised project(s) in year *t*.

Information on balancing quality is reported in the on-line Appendix for a selection of samples, including the full sample, the sample for the global financial crisis period (2008-2010), firms in the first quartile of the employment distribution, firms in third quartile of the age distribution, and firms in the fourth quartile of the market-share distribution.[12] As a general rule, we sought covariate balance for two moments (mean and variance) and evaluated the quality of the covariate balance using the standardised difference between means of treated and reweighted control groups.

We are aware of the ongoing debate on the metrics for evaluating covariate balance quality. Two most used measures are the standardized difference (bias) in means and t-tests for differences in means (Rosenbaum and Rubin 1983). Imai, King, and Stuart (2008) criticize the

[12] 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.

use of t-tests and argue for QQ plot summary statistics. A recent simulation study (Franklin et al., 2013) indicates that several metrics are appropriate, including the standardized difference in means. The latter also has the advantage that it automatically evaluates balance on all covariates simultaneously and can incorporate balance on interactions among covariates. Furthermore, simulation results in Hainmueller (2012) indicate that EB weighting yields standardised differences very close to zero and p-values of almost 1, when t-tests are conducted for mean differences. Therefore, we rely on standardised difference as the metric for evaluating covariate balance.

The standardised difference between means of the subsidised and unsubsidised firms are reported in column 10 (before EB weighting) and column 11 (after EB weighting) of each table in the online Appendix. Not surprisingly, the standardised mean differences are large before EB weighting. Hence, effect size estimates without weighting or matching would be evidently biased. In contrast, the standardised differences after EB weighting are negligible and any difference from zero is usually observed only after the third decimal point. A comparison of the covariate variances between subsidised and non-subsidised firms after EB weighting also indicates that the variances are usually equal up to three decimal points.[13]

Table 5 reports ATT estimates for different R&D input types over the full estimation period 1998-2012 (column 1) and during two subperiods that correspond to the *dot.com* crisis and global financial crisis (columns 2 and 3). The first finding to note is that the ATT is positive and statistically and practically significant (around 2% to 5%) for privately funded R&D, R&D personnel intensity, development R&D and experimental R&D. However, the ATTs for basic R&D, capital R&D and extra-mural R&D are much smaller: they are practically insignificant albeit statistically significant. The ATT for privately funded R&D indicates additionality, whereas the rest indicate absence of full crowding out. The second point to note is that the ATTs are always smaller or insignificant during crisis periods (columns 2 and 3).

These findings in Table 5 provide empirical support for **H1**, which stipulates that R&D subsidies are less effective in generating additionality effects during financial crises or when the investment is in basic research. Comparing the rows, we observe that the subsidy's effect

[13] Because of the high dimensionality of the selection process (hence, the EB) with 139 covariates described in Table A2 in the Appendix, we were unable to obtain covariate balance in terms of skewness.

on basic R&D (row 3) is smaller than the effect on all other R&D types, except extra-mural and capital R&D in rows 6 and 7. Comparing columns, we see that the ATTs during crisis periods (columns 2 and 3) are either insignificant or consistently smaller than the ATTs for the full sample.[14]

Table 5. UK R&D subsidy and R&D investment during crisis periods: ATTs for different R&D types.

Subsidy effects on growth of:	(1) Full sample	(2) dot-com crisis 2000-2002	(3) Global financial crisis 2008-2010
1. Private R&D intensity	.0457***	.0217	.0235***
	(.0060)	(.0167)	(.0075)
2. R&D personnel intensity	.0456***	.0129	.0365***
	(.0066)	(.0151)	(.0111)
3. Basic R&D intensity	.0063***	.0113***	.0019***
	(.0015)	(.0040)	(.0005)
4. Experimental R&D intensity	.0158***	.0011	.0214***
	(.0072)	(.0104)	(.0052)
5. Applied R&D intensity	.0244***	.0153*	.0077
	(.0036)	(.0092)	(.0079)
6. Extramural R&D intensity	.0027***	.0014	.0011
	(.0006)	(.0021)	(.0015)
7. Capital R&D intensity	.0040***	.0035*	.0024**
	(.0007)	(.0018)	(.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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The smaller subsidy effects on basic R&D or during crisis periods are due to higher return uncertainty, which reduces the risk-averse firm's optimal level of R&D investment. Returns to basic research are more uncertain because the its immediate commercial applications are less clear and even successful basic research outcomes are converted into profitable product lines with long time lags.[15] R&D return uncertainty is also higher during crisis periods because of

[14] 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).

[15] The ONS survey defines basic research as "work undertaken primarily for the advancement of scientific knowledge without a specific practical application in view". This definition is based on the Frascati Manual, accessible at https://www.oecd.org/sti/frascati-manual-2015-9789264239012-en.htm

higher product-market uncertainties, which increase the volatility of the returns on all types of R&D investment. Therefore, the responsiveness of risk-averse firms to the subsidy is weaker if the investment is basic R&D or undertaken during crisis periods. Our findings are in line with: (i) empirical findings in Aristei et al., (2017) who report weaker additionality effects during the global financial crisis; (ii) predictions from investment models of R&D under uncertainty (Czarnitzki and Toole, 2013); and (iii) predictions from 'waiting' models where R&D investment is persistent and less responsive to policy interventions due to higher return uncertainty (Bloom, 2007).

Results in Table 5 also lends support (or is congruent with) the assumption of complementarity between different R&D types. Although the specific R&D type targeted by the subsidy is not known, the effect of the subsidy on aggregate and individual measures of R&D investment is in the same direction. This is particularly important in the case of extra-mural (contracted-out) R&D that, by definition, is not eligible for R&D subsides. Although the ATTs for extra-mural R&D are small, they indicate that firms that increase intra-mural R&D investment in response to the subsidy also tend to increase their contracted-out R&D investment. This finding resonates with Lokshin et al., (2008), who report complementarity between the productivity effects of intramural and extra-mural R&D investment.

In Table 6 we provide evidence on how the subsidy's effects on basic R&D vary by firm type in terms of their locations in the age, size and market share distributions. The results indicate that the small and positive subsidy effects on the growth of basic R&D intensity decline and eventually become insignificant among older and larger firms and firms with larger market shares. This finding indicates that the subsidy's effect is heterogeneous not only between R&D types but also between different firm cohorts even in the case of basic research. The overall effect reported in Table 5 above for the full sample is mainly due to the response of basic R&D investment in smaller and younger firms, which receive a small fraction (around 8-10 percent) of the subsidy allocations.

Results in Tables 5 and 6 indicate that the 'financing constraint' argument is necessary but not sufficient to explain the subsidy's effect on business R&D investment in general or basic R&D investment in particular. It is true that financial constraints are more biting when the investment is in basic research due to higher levels of information asymmetry between the firm and private funders; or during recessions due to pro-cyclical behaviour of the capital markets. That is why several studies (e.g., Czarnitzki and Toole, 2007; Hall, 1992; 2002; 2010; Minton and Schrand, 1999) recommend direct or indirect public support for basic R&D and R&D investment during crisis periods. This is also reflected in the R&D support policies in the UK and elsewhere, which are more generous in the case of basic research or during crisis periods. Our proposed analytical framework, however, can and does explain the results through risk aversion. R&D subsidies do alleviate the financing constraint in the case of basic research or during crisis periods, but this positive effect is mitigated by the adverse effects of higher rates with which firms discount the future returns on riskier projects.

Table 6: Effects of UK R&D subsidy on basic R&D by firm-type quartiles: ATTs for growth of basic R&D intensity.

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

Notes: The sample size for different R&D types differs because firms do undertake all types of R&D investment every year. N_1 and N_0 are numbers of treated and control observations. For other notes, see Table 6 and 7.

Combining the findings in Tables 5 and 6 with descriptive evidence in Tables 1 – 4, we can see that the effect of the subsidy on business R&D investment tends to be smaller or insignificant when the subsidy rate is higher. 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 support for business R&D as follows: *it is socially optimal 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 7 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 assumes 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 firms the optimal level of R&D intensity among firms closer to the R&D frontier would be higher and their R&D gaps would be smaller even in the absence of the subsidy, as indicated in equation (2). Therefore, the subsidy would induce smaller and eventually insignificant additional R&D investment as the firms move closer to the R&D frontier.

Table 7. Effects of UK R&D subsidy by R&D type and firm proximity to R&D frontier: ATTs for growth of different R&D investment intensities.

	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 N_0= 2782 N_1= 22173	-.0007 (.0084)	.0085 (.0077)	-.0009 (.0070)	.0007 (.0029)	-.0016 (.0025)	-.0017 (.0023)	.0013 (.0009)
Quartile 2 N_0= 1301 N_1= 23655	.0037 (.0053)	.0068 (.0051)	.0027 (.0042)	.0018 (.0023)	.0003 (.0005)	.0004 (.0009)	.0001 (.0004)
Quartile 3 N_0= 485 N_1= 24470	.0248*** (.0055)	.0249*** (.0055)	.0116*** (.0037)	.0112*** (.0033)	.0018 (.0015)	.0033** (.0016)	.0038*** (.0009)
Quartile 4 N_0= 470 N_1= 24486	.0495*** (.0134)	.0611*** (.0158)	.0271*** (.0057)	.0263*** (.0099)	.0134*** (.0031)	.0051*** (.0009)	.0076*** (.0014)

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. The number of subsidised and control firm-year observations varies between quartiles as firms are not sampled every year in the ONS survey. Each quarter has the same number of non-missing common observations. N_1 and N_0 are the numbers of treated and control observations respectively.

In the data (Table 4), the subsidised firms have higher levels of R&D intensity compared to unsubsidised sample. As such, the subsidised sample includes a higher proportion of firms closer to the R&D frontier. Thus, the combination of the descriptive evidence and ATT estimations reveal a second *conundrum* in public support for business R&D: *it is socially optimal 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*. Our proposed theoretical framework explains this conundrum through insights form the theory of contracts (Laffont and Mortimort, 2009; Akcigit et al., 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 also provides a unified theoretical underpinning for similar findings reported in Lach (2002), Lee (2011), and Wanzenböck et al., (2013).

The estimated parameters in panel A (Table 8) indicate that the ATTs are declining with age, size and market share for two aggregate measures R&D inputs: privately funded R&D intensity and R&D personnel intensity. A similar pattern is evident in panel B, where we report ATTs for two specific measures: extramural R&D and capital R&D investment. In both panels, the ATTs are the largest among youngest and smallest firms and for companies in the most competitive markets (quartile 1). Then the ATTs decline and become insignificant as firms grow in age or size or capture larger market shares. The pattern clearly indicates the positive effect of the subsidy on these R&D inputs in the full sample (Table 5) conceals a high degree of heterogeneity as the full-sample effect is driven by the impact among smaller/younger firms. As such, the sample average ATT is a poor basis for evidence-based public policy.

Results in Table 8 lend support to **H3**, which postulates that the subsidy's effect on business R&D investment is negatively related to firm age, size, and market share. This is because the age, size and market share of the firm are determined by the efficiency with which it converts the R&D investment into innovative product lines. Stated differently, firms with higher R&D productivity are older and larger in accordance with Schumpeterian models of innovation (Aghion et al., 2014; 2015). Also, firms with higher R&D productivity are those with narrower R&D gaps and with lower levels of responsiveness to R&D subsidies as suggested by the theory of contracts (Laffont and Mortimort, 2009; Akcigit et al., 2019). Finally, the findings in Table 8 are consistent with Lach (2002), Gonzalez and Pazo (2008), and Wanzenböck et al., (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.

Table 8: Effects of UK R&D subsidy by age, size and market share quartiles.

Panel A: ATTs for 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		
	age	employment	deflated turnover	Age	employment	deflated turnover
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: ATTs for extramural R&D and capital R&D intensities.

	Growth of extramural R&D intensity			Growth of capital R&D expenditures intensity		
	age	employment	deflated turnover	age	employment	deflated turnover
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; employment is number of employees, turnover is deflated with GDP deflator, market share is estimated at 3-digit SIC industry level (254 industries). N_1 and N_0 are numbers of treated and control observations. For other notes, see Table 6 and 7.

The results in Table 8 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, but such firms are less likely to produce additionality effects*. This is particularly the case in the UK, where disproportionately higher percentages of the subsidy are allocated to older and larger firms.

Our final remark relates to the robustness of the findings discussed above to a wide range of sensitivity checks, reported in the on-line Appendix. 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 different findings. Then we restricted the sample to manufacturing firms only (Table A4) and used initial weights from a coarsened exact matching (CEM) procedure. Finally, we regressed the estimated ATTs for private R&D intensity growth in the full sample on firm characteristics and crisis periods. The results from all sensitivity checks are consistent with the results reported and discussed above. Therefore, we are confident that our main findings are not driven by sample selection or initial weights used by the entropy balancing routine.

6. Conclusions

In this paper, we have evaluated the effects of R&D subsidies on separate and aggregate measures business R&D investment theoretically and empirically. Theoretically, we have drawn on the theory of contracts and Schumpeterian models of innovation to derive 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 unobserved R&D productivity, R&D gap and risk aversion moderate the effect of the subsidy on business R&D investment; and (ii) mapping the unobserved moderating 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 etc.).

Our second contribution consists of enhancing the evidence base with findings that are consistent with the proposed hypotheses and remain robust to a wide range of sensitivity checks. The evidence indicates that the effects of R&D subsidies: (i) are positive in the full sample, but they conceal a high degree of heterogeneity; (ii) are smaller or insignificant when investment is in basic research or during crisis periods; and (iii) decline and eventually become insignificant as firms are older, larger and closer to the R&D frontier. Our work strengthens the case for paying attention to heterogeneity in the subsidy's effects on business R&D and provides a unifying theoretical framework that explains the sources of observed heterogeneity.

Our third contribution is to place the policy conundrums in public support for business R&D under the spotlights. Specifically, we have demonstrated that the subsidy regime would deliver sub-optimal subsidy allocations and second-best firm R&D effort due to information asymmetry and risk aversion. Under information asymmetry, subsidy allocations and the firm's R&D effort remain suboptimal due a trade-off between the funder's attempts at ensuring efficiency by satisfying the participation constraints of the high-R&D-productivity firms and the latter's ability to extract informational rents by concealing their true types. These second-best outcomes are exacerbated when firms are more risk averse. The main policy-relevant take-away from our findings can be stated as follows: it is socially optimal to subsidise business R&D when the R&D gap is wider or the subsidised firms are more likely to convert R&D inputs into innovative product lines, but the subsidy would be less effective under these circumstances. This is due to risk aversion in the former and informational rent extraction in the latter case.

Our work also expands the range of treatment-effect estimators used to evaluate the subsidy's effect on business R&D investment. The entropy balancing (EB) method allows for better covariate balancing for a larger number of covariates; has been shown to perform better than propensity-score matching routines; is comparable with the doubly-robust estimator when both the selection and outcome models in the latter are correctly specified; and produces smaller bias than conventional doubly robust estimators. Furthermore, the EB method has the advantage of making full use of the information in the control sample.

Our findings suggest that R&D subsidies may be 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. To bridge the R&D gap, direct support for R&D should target small and young firms where it is more likely to generate additionality effects. To maximise social welfare despite the efficiency shortfall from targeting small and young firms, the policy can rely on *ex post* 'innovation prizes' that reward successful innovations by all innovators. Nevertheless, *ex post* innovation rewards require further research on measuring and verifying innovation quality; and more transparent innovation systems in which private innovators and public funders interact.

References

Abel, A., Dixit, K., Eberly, J. and Pindyck, R. (1996). 'Options, the value of capital, and investments', *Quarterly Journal of Economics*, vol. 111(3), pp. 753-777.

Abel, A. and Eberly, J. (1996). 'Optimal investment with costly reversibility', *Review of Economic Studies*, vol. 63(4), pp. 581-593.

Acemoglu, D. and Linn, J. (2004). 'Market size in innovation: Theory and evidence from the pharmaceutical industry', *Quarterly Journal of Economics*, vol. 119(3), pp. 1049-1090.

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

Akcigit, U., Hanley, D. and Stantcheva, S. (2019). 'Optimal taxation and R&D policies', National Bureau of Economic Research Working Paper No. w22908, Cambridge, MA.

Akcigit, U. and Kerr, W. (2018). 'Growth through heterogeneous innovations', *Journal of Political Economy*, vol. 126(4), pp. 1374-1443.

Aerts K. and Schmidt, T. (2008). 'Two for the price of one? Additionality effects of R&D subsidies: A comparison between Flanders and Germany', *Research Policy*, vol. 37, pp. 806-822.

Aghion, P., Askenazy, P., Berman, N., Cette, G. and Eymard, L. (2012). 'Credit constraints and the cyclicality of R&D investment: Evidence from France'. *Journal of the European Economic Association*, vol. 10(5), pp. 1001-1024.

Aghion, P., Akcigit, U. and Howitt, P. (2014). 'What do we learn from Schumpeterian growth theory?' (in Aghion, P., Durlauf, S., eds.), *Handbook of Economic Growth*, vol. 2, pp. 515-563, Elsevier.

Aghion, P., Akcigit, U., and Howitt, P. (2015). 'Lessons from Schumpeterian growth theory'. *American Economic Review*, vol. 105(5), pp. 94-99.

Amusa, L., Zewotir, T. and North, D. (2019). 'Examination of entropy balancing technique for estimating some standard measures of treatment effects: a simulation study'. *Electronic Journal of Applied Statistical Analysis*, vol. 12(2), pp. 491-507.

Aristei, D., Sterlacchini, A. and Venturini, F. (2017). 'Effectiveness of R&D subsidies during the crisis: firm-level evidence across EU countries'. *Economics of Innovation and New Technology*, vol. 26(6), pp. 554-573.

Arrow, K. (1962). 'Economic welfare and the allocation of resources for invention', (in Nelson R.R., ed.), The Rate and Direction of Inventive Activity: Economic and Social Factors, National Bureau of Economic Research, Conference Series, Princeton University Press, Princeton, pp. 609-625.

Arrow, K. (1996). 'The economics of information: An exposition'. *Empirica*, vol. 23(2), pp. 119-128.

Baron, D. and Myerson, R. (1982). 'Regulating a monopolist with unknown costs', *Econometrica*, vol. 50(4), pp. 911-930.

Bernanke, B. (1983). 'Irreversibility, uncertainty, and cyclical investment', *Quarterly Journal of Economics*, vol. 98(1), pp. 85-106.

Blackwell, M., Iacus, S., King, G., and Porro, G. (2009). 'Cem: Coarsened exact matching in Stata', *The Stata Journal*, vol. 9(4), pp. 524-546.

Blank, D. and Stigler, G. (1957). *The demand and supply of scientific personnel*, New York: National Bureau of Economic Research.

Bloom, N. (2007). 'Uncertainty and the dynamics of R&D', *American Economic Review*, vol. 97(2), pp. 250-255.

Bloom, N., Bond, S. and van Reenen, J. (2007). 'Uncertainty and investment dynamics', *Review of Economic Studies*, vol. 74(2), pp. 391-415.

Bogliacino, F. and Pianta, M. (2010). 'Innovation and employment: A reinvestigation using revised Pavitt classes', *Research Policy*, vol. 39(6), pp. 799-809.

Bolton, P. and Dewatripont, M. (2005). *Contract theory*, London: MIT Press.

Bouvatier, V. and Lepetit, L. (2008). 'Banks' procyclical behavior: Does provisioning matter?' *Journal of International Financial Markets, Institutions and Money*, vol. 18(5), pp. 513-526.

Bose, A., Pal, D. and Sappington, D. (2011). 'On the performance of linear contracts', *Journal of Economics & Management Strategy*, vol. 20(1), pp. 159-193.

Branstetter, L. and Sakakibara, M. (1998). 'Japanese research consortia: A microeconometric analysis of industrial policy.' *Journal of Industrial Economics*, vol. 46(2) , 207–233.

Caggese, A. (2012). 'Entrepreneurial risk, investment, and innovation'. *Journal of Financial Economics*, vol. 106(2), pp. 287-307.

Carroll, G. (2015). 'Robustness and linear contracts', *The American Economic Review*, vol. 105(2), pp. 536-563.

Chari, V., Golosov, M. and Tsyvinski, A. (2012). 'Prizes and patents: Using market signals to provide incentives for innovations', *Journal of Economic Theory*, vol. 147(2), pp. 781-801.

Chassang, S. (2013). 'Calibrated incentive contracts', *Econometrica*, vol. 81(5), pp. 1935-1971.
Clausen, T. (2009). 'Do subsidies have positive impacts on R&D and innovation activities at the firm level?', *Structural Change and Economic Dynamics*, vol. 20(4), pp. 239-253.

Cohen, W. and Levinthal, D. (1989). 'Innovation and learning: The two faces of R&D.' *Economic Journal*, vol. 99(397), 569–596.

Czarnitzki, D. and Hottenrott, H. (2011). 'R&D investment and financing constraints of small and medium-sized firms', *Small Business Economics*, vol. 36(1), pp. 65-83.

Czarnitzki, D., Hottenrott, H. and Thorwarth, S. (2011). 'Industrial research versus development investment: the implications of financial constraints', *Cambridge Journal of Economics*, vol. 35(3), pp. 527–544.

Czarnitzki, D. and Hussinger, K. (2004). 'The link between R&D subsidies, R&D spending and technological performance'. ZEW Centre for European Economic Research Discussion Paper 04-056.

Czarnitzki, D. and Lopes-Bento, C. (2013). 'Value for money? New microeconometric evidence on public R&D grants in Flanders', *Research Policy*, vol. 42(1), pp. 76-89.

Czarnitzki, D. and Toole, A. (2007). Business R&D and the interplay of R&D subsidies and product market uncertainty. *Review of Industrial Organization*, vol. 31(3), 169-181.

Czarnitzki, D., Toole, A., 2013. 'The R&D investment–uncertainty relationship: Do strategic rivalry and firm size matter?', *Managerial and Decision Economics*, vol. 34(1), pp. 15-28.

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

Deschryvere, M. (2014). 'R&D, firm growth and the role of innovation persistence: An analysis of Finnish SMEs and large firms', *Small Business Economics,* vol. 43(4), pp. 767-785.

Dimos, C. and Pugh, G. (2016). 'The effectiveness of R&D subsidies: A meta-regression analysis of the evaluation literature', *Research Policy,* vol. 45(4), pp. 797-815.

Dorie, V., Hill, J., Shalit, U., Scott, M., and Cervone, D. (2019). 'Automated versus do-it-yourself methods for causal inference: Lessons learned from a data analysis competition', *Statistical Science,* vol. 34(1), pp. 43-68.

Dubois, P., De Mouzon, O., Scott-Morton, F. and Seabright, P. (2015). 'Market size and pharmaceutical innovation', *The RAND Journal of Economics*, vol. 46(4), pp. 844-871.

Duguet, E., 2004. 'Are R&D subsidies a substitute or a complement to privately funded R&D? Evidence from France using propensity score methods for non-experimental data', *Revue d'Economie Politique,* vol. 114(2), pp. 263-292.

Fabrizio, K.R. and Tsolmon, U. (2014). 'An empirical examination of the procyclicality of R&D investment and innovation', *Review of Economics and Statistics,* vol. 96(4), pp. 662-675.

García-Quevedo, J. (2004). 'Do public subsidies complement business R&D? A meta-analysis of the econometric evidence', *Kyklos*, vol. 57(1), pp. 87-102.

Geroski, P. (1995a). 'What do we know about entry?' *International Journal of Industrial Organization*, vol. 13(4), pp. 421-440.

Geroski, P.A. (1995b). 'Do spillovers undermine the incentive to innovate?' (in Dowrick, S. Ed.., ed.), Economic Approaches to Innovation, Edwar Elgar, Aldershot, pp. 76–97.

Ghiglino, C. and Tabasso, N. (2016). 'Risk aversion in a model of endogenous growth'. *Journal of Mathematical Economics*, vol. 64, pp. 30-40.

Goel, R. K. and Ram, R. (2001). 'Irreversibility of R&D investment and the adverse effect of uncertainty: Evidence from the OECD countries', *Economics Letters,* vol. 71(2), pp. 287-291.

Griliches, Z. (1979). 'Issues in assessing the contribution of research and development to productivity growth', *Bell Journal of Economics*, vol. 10(1), pp. 92–116.

Hainmueller, J. and Xu, Y. (2013). 'Ebalance: A Stata package for entropy balancing', *Journal of Statistical Software*, vol. 54(7), pp. 1-18.

Hainmueller, J. (2012). 'Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies', *Political Analysis*, vol. 20(1), pp. 25-46.

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*, vol. 18, pp. 35-51.

Hall, B. (2010). 'The financing of innovative firms', *Review of Economics and Institutions*, vol. 1(1), pp. 1–30.

Haltiwanger, J., Jarmin, R. and Miranda, J. (2013). 'Who creates jobs? Small versus large versus young', *Review of Economics and Statistics*, vol. *95*(2), pp. 347-361.

Heckman, J., Ichimura, H. and Todd, P. (1998). 'Matching as an econometric evaluation estimator', *Review of Economic Studies*, vol. 65(2), pp. 261-294.

Hirano, K., Imbens, G.W., Rubin, D. and Zhou, X. (2000). 'Assessing the effect of an influenza vaccine in an encouragement design', *Biostatistics*, vol. 1(1), pp. 69-88.

Holmstrom, B. and Milgrom, P. (1991). 'Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design', *Journal of Law, Economics, and Organization*, vol 7, pp. 24-52.

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

Hussinger, K. (2008). 'R&D and subsidies at the firm level: An Application of parametric and semi-parametric two-step selection models', *Journal of Applied Econometrics*, vol. 23, pp. 729-747.

Imai, K., King, G. and Stuart, E.A. (2008). 'Misunderstandings between experimentalists and observationalists about causal inference', *Journal of the Royal Statistical Society: Series A (statistics in society)*, vol. 171(2), pp. 481-502.

King, G. and Nielsen, R. (2019). 'Why propensity scores should not be used for matching', *Political Analysis*, vol. 27(4), pp. 1-20.

Kravitz, R., Duan, N., and Braslow, J. (2004). 'Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages', *The Milbank Quarterly*, vol. 82(4), pp. 661-687.

Kullback, S. (1959). *Information Theory and Statistics*, New York: Wiley.

Kwon, I. (2010). 'R&D portfolio and market structure.' *Economic Journal*, vol. 120(543), pp. 313-323.

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

Laffont, J. and Martimort, D. (2002). *The Theory of Incentives: The Principal-agent Model*, NJ: Princeton University Press.

Laffont, J. and Tirole, J. (1986). 'Using cost observation to regulate firms', *Journal of Political Economy*, vol. 94(3, Part 1), pp. 614-641.

Lee, C.-Y. (2011). 'The differential effects of public R&D support on firm R&D: Theory and evidence from multi-country data', *Technovation*, vol. 31, pp. 256-269.

Lokshin, B., Belderbos, R. and Carree, M. (2008). ‘The productivity effects of internal and external R&D: Evidence from a dynamic panel data model’, *Oxford Bulletin of Economics and Statistics*, vol. 70 (3), pp. 399–413.

Marino, M., Lhuillery, S., Parrotta, P. and Sala, D. (2016). ‘Additionality or crowding-out? An overall evaluation of public R&D subsidy on private R&D expenditure’, *Research Policy,* vol. 45(9), pp. 1715-1730.

Marcus, J. (2013). ‘The effect of unemployment on the mental health of spouses–Evidence from plant closures in Germany’, *Journal of Health Economics*, vol. 32(3), pp. 546-558.

McMullin, J. and Schonberger, B. (2020). ‘Entropy-balanced accruals’, *Review of Accounting Studies*, vol. 25, pp. 84–119.

Meuleman, M. and de Maeseneire, W. (2012). ‘Do R&D subsidies affect SMEs’ access to external financing?’, *Research Policy*, vol. 41(3), pp. 580-591.

Minton, B. and Schrand, C. (1999). ‘The impact of cash flow volatility on discretionary investment and the cost of debt and equity financing’, *Journal of Financial Economics*, vol. 54, pp. 423-460.

Mantovani, A. (2006). ‘Complementarity between product and process innovation in a monopoly setting.’ *Economics of Innovation and New Technology*, vol. 15(03), pp. 219-234.

Mohnen, P., Polder, M. and van Leeuwen, G. (2018). “ICT, R&D and organizational innovation: Exploring complementarities in investment and production” National Bureau of Economic Research (NBER) Working Papers, no. 25044.

Nelson, R. (1959). ‘The simple economics of basic scientific research’, *Journal of Political Economy*, vol. 67(3), pp. 297-306.

Neuenkirch, M. and Neumeier, F. (2016). ‘The impact of US sanctions on poverty’, *Journal of Development Economics*, vol. 121, pp. 110-119.

Nilsen, Ø., Raknerud, A., and Iancu, D. (2018). ‘Public R&D support and firms' performance: A panel data study’, NHH Dept. of Economics Discussion Paper No. 13/2018.

Office for National Statistics (2012). Coverage of the Business Enterprise Research & Development Survey, Office of National Statistics, http://www.ons.gov.uk/ons/guide-method/method-quality/specific/business-and-energy/research---development-surveys/information-note--coverage-of-the-business-enterprise-r-d-survey.pdf.

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

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

Packham, N. (2018). ‘Optimal contracts under competition when uncertainty from adverse selection and moral hazard are present’, *Statistics & Probability Letters*, vol. 137, pp. 99-104.
Pavitt, K. (1984). ‘Sectoral patterns of technical change: towards a taxonomy and a theory’, *Research Policy,* vol. 13(6), pp. 343-373.

Raymond, W., Mohnen, P., Palm, F., and van Der Loeff, S. (2010). ‘Persistence of innovation in Dutch manufacturing: Is it spurious?’, *Review of Economics and Statistics*, vol. 92(3), pp. 495-504.

Romer, P. (1990). 'Endogenous technological change', *Journal of Political Economy,* vol. 98(5), pp. 71–102.

Rosenbaum, P. and Rubin, D. (1983). 'The central role of the propensity score in observational studies for causal effects', *Biometrica,* vol. 70, pp. 41-55.

Rubin, D. (2001). 'Using propensity scores to help design observational studies: application to the tobacco litigation', *Health Services & Outcomes Research Methodology*, vol. 2, pp. 169–188.

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

Sappington, D. (1982). 'Optimal regulation of research and development under imperfect information', *The Bell Journal of Economics,* vol. 13(2), pp. 354-368.

Smith, J. and Todd, P. (2005). 'Does matching overcome LaLonde's critique of nonexperimental estimators?', *Journal of Econometrics,* vol. 125(1-2), pp. 305-353.

Takalo, T., Tanayama, T. and Toivanen, O. (2013). 'Estimating the benefits of targeted R&D subsidies', *Review of Economics and Statistics*, vol. 95(1), pp. 255-272.

Triguero, A. and Córcoles, D. (2013). 'Understanding innovation: An analysis of persistence for Spanish manufacturing firms', *Research Policy,* vol. 42(2), pp. 340-352.

Ugur, M., Trushin, E., and Solomon, E. (2016a). 'Inverted-U relationship between R&D intensity and survival: Evidence on scale and complementarity effects in UK data', *Research Policy,* vol. 45(7), pp. 1474-1492.

Ugur, M., Trushin, E. and Solomon, E. (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., Scherngell, T. and Fischer, M. (2013). 'How do firm characteristics affect behavioural additionalities of public R&D subsidies? Evidence for the Austrian transport sector', *Technovation*, vol. 33(2-3), pp. 66-77.

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

Zhao, Q. and Percival, D. (2017). 'Entropy balancing is doubly robust', *Journal of Causal Inference*, vol. 5(1), pp. 1-19.

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

Information asymmetry, risk aversion and R&D subsidies: Effect-size heterogeneity and policy conundrums
ONLINE APPENDIX.

This on-line appendix provides summary statistics, sensitivity checks and entropy balancing quality diagnostics for the paper above. The tables here are referred to in the main paper.

Table A1. Summary statistics for one-year-lagged covariates in the entropy balance (EB) weighting model.

	Control group (N_0 = 10282)		Treated group (N_1 = 133563)	
	Mean	Std. dev.	Mean	Std. dev.
Log employment	4.427	2.113	3.112	1.729
UK ownership dummy	0.759	0.428	0.870	0.336
R&D intensity	0.071	0.155	0.089	0.147
Log age	2.757	0.724	2.597	0.795
Live local units (plants)	10.970	136.007	2.917	57.540
Log deflated turnover	8.965	2.700	7.301	2.343
Log deflated turnover per employee	4.617	1.171	4.350	0.952
Share of R&D personnel in empoyment	0.081	0.157	0.088	0.116
Def. turnover growth	0.076	0.622	0.052	0.551
Market share	0.026	0.086	0.006	0.036
UK subsidy dummy	0.418	0.493	0.967	0.178
UK and EU sub. dummy	0.280	0.449	0.787	0.410
Civil R&D only dummy	0.753	0.431	0.424	0.494
Tax credit for small firms	0.505	0.500	0.819	0.385
Tax credit for large firms	0.215	0.411	0.064	0.245
Start-up firm dummy	0.044	0.205	0.071	0.258
Young firm dummy	0.150	0.357	0.204	0.403
SME dummy	0.502	0.500	0.535	0.499
Private R&D intensity	0.070	0.154	0.079	0.130
Basic R&D intensity	0.005	0.037	0.009	0.027
Experimental R&D intensity	0.027	0.083	0.030	0.053
Extramural R&D expenditure intensity	0.006	0.032	0.005	0.019
Capital R&D expenditure intensity	0.003	0.012	0.005	0.009
Indicator for SMEs R&D tax credit change in 2008	0.003	0.012	0.414	0.493
Indicator for large firm R&D tax credit change in 2008	0.158	0.364	0.013	0.111
Indicator that SMEs cannot use R&D subsidies for tax deduction	0.055	0.228	0.824	0.381
Applied R&D intensity	0.027	0.085	0.038	0.075
Herfindahl index	0.123	0.138	0.094	0.103
Crisis dummy 2008-2010	0.308	0.462	0.280	0.449
Dot.com crisis dummy 2000-2002	0.166	0.372	0.116	0.320
Pavitt class1 dummy	0.162	0.368	0.322	0.467
Pavitt class2 dummy	0.096	0.295	0.229	0.420
Pavitt class3 dummy	0.156	0.363	0.106	0.308
Pavitt class4 dummy	0.461	0.499	0.293	0.455

Notes: The EB weighting model consists of 133 covariates, 17 pre-treatment covariates at the firm level; 7 covariates at the industry level; 6 dummy variables that capture the receipt of EU subsidies and the change in the UK's R&D tax credit regime; 2 crisis dummies for 2000-2002 and for 2008-2010; 15 year dummies; and 84 industry dummies at the two-digit SIC level. Year and industry dummies are not reported here to save space. All variables except crisis dummies and R&D tax regime dummies are lagged one year. Market share and the Herfindahl index are calculated within 3-digit industries. The full sample excludes firm/year observations with privately funded R&D intensity greater than 1. Minimum and maximum values are not reported to comply with non-disclosure requirements of the data host, UK Data Service. Pavitt technology classes are from Pavitt (1984), as revised slightly by Bogliacino and Pianta (2010).

Table A2. Sensitivity check 1- ATTs using logarithm of private R&D intensity: By age, size and market share quartiles.

	By age quartiles	By employment quartiles	By deflated turnover quartiles	By market share quartiles
Quartile 1	.0664***	.0922***	.1055***	0.1181***
	(.0121)	(.0206)	(.0211)	(0.0229)
	N =36,022	*N = 36,022*	*N = 36,022*	*N = 36,022*
Quartile 2	.0368**	.0265***	.0324***	0.0225**
	(.0145)	(.00837)	(.0083)	(0.0096)
	N = 36,073	*N = 36,073*	*N = 36,073*	*N = 36,073*
Quartile 3	.0159	.0139	.0106	0.0051
	(.0163)	(.0098)	(.0077)	(0.0049)
	N = 36,073	*N = 35,570*	*N = 35,570*	*N = 35,570*
Quartile 4	.0009	-.0131	-.0114	-0.0053
	(.0007)	(.0143)	(.0097)	(0.0055)
	N = 34,109	*N = 34,109*	*N = 34,109*	*N = 34,109*

Notes: The outcome variable is the logarithm of R&D input intensity. Bootstrap standard errors are estimated with 100 iterations. *N* is number of firm-year observations in the quartile. The number of firm-year observations varies between quartiles as firms are not sampled every year in the ONS survey. Market share is the firm's share in industry turnover within 254 SIC 3-difit industries. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3a. Sensitivity check: ATTs for surviving firms only for R&D input types by employment quartiles.

Employment quartiles:	Growth of private R&D intensity	Growth of R&D personnel intensity	Growth of basic R&D intensity	Growth of experimental R&D intensity	Growth of applied R&D intensity	Growth of extramural R&D intensity	Growth of capital R&D intensity
Quartile 1	.0706***	.0518***	.0052***	.0253**	.0482***	.0027	.0067**
N_0= 749	(.0208)	(.0135)	(.0016)	(.0117)	(.0140)	(.0022)	(.0025)
N_1 = 27296							
Quartile 2	.0134	.0289**	.0030**	.0036	.0012	.0077	.0039***
N_0= 1432	(.0127)	(.0146)	(.0014)	(.0068)	(.0006)	(.0054)	(.0009)
N_1 = 26836							
Quartile 3	.0002	-.0047	-.0004	.0010	-.0019	.0019	.0023***
N_0= 1661	(.0072)	(.0126)	(.0008)	(.0055)	(.0029)	(.0070)	(.0006)
N_1 = 25758							
Quartile 4	.0000	.0009	-.0007	-.0001	.0009	.0006	.0003
N_0= 4006	(.0024)	(.0040)	(.0006)	(.0022)	(.0012)	(.0005)	(.0002)
N_1 = 23824							

Note: N_0 - number of control observations, N_1 - number of treated observations.

Table A3b. Sensitivity check: ATTs for surviving firms only for R&D input types by deflated turnover quartiles.

Deflated turnover quartiles:	Growth of private R&D intensity	Growth of R&D personnel intensity	Growth of basic R&D intensity	Growth of experimental R&D intensity	Growth of applied R&D intensity	Growth of extramural R&D intensity	Growth of capital R&D intensity
Quartile 1	.0655***	.0429***	.0059***	.0227***	.0466***	.0031	.0074***
N_0= 813	(.0197)	(.0137)	(.0017)	(.0088)	(.0155)	(.0021)	(.0018)
N_1 = 27078							
Quartile 2	.0211*	.0242**	.0020	.0091	.0094	.0022	.0034**
N_0= 1484	(.0127)	(.0111)	(.0012)	(.0056)	(.0060)	(.0015)	(.0014)
N_1 = 26406							
Quartile 3	.0037	.0097	-.0001	.0011	.0015	.0013	.0013**
N_0= 1704	(.0057)	(.0091)	(.0011)	(.0037)	(.0024)	(.0017)	(.0005)
N_1 = 26187							
Quartile 4	-.0010	.0055	-.0004	-.0002	.0005	.0007	-.0001
N_0= 3847	(.0020)	(.0037)	(.0003)	(.0016)	(.0009)	(.0005)	(.0004)
N_1 = 24043							

Note: N_0 - number of control observations, N_1 - number of treated observations.

Table A4. Sensitivity check: ATTs for manufacturing firms, using two-year-lagged covariates, or R&D intensity cut-off point at 99th percentile instead of 96th percentile.

	Manufacturing firms only	All EB covariates lagged two years	R&D intensity cut-off point at 99 percentile
Private R&D intensity	.0097	.0325***	.0334***
	(.0069)	(.0082)	(.0081)
R&D personnel intensity	.0146**	.0332***	.0312***
	(.0064)	(.0062)	(.0059)
Basic R&D intensity	.0019**	.0044**	.0048**
	(.0007)	(.0019)	(.0019)
Experimental R&D intensity	.0046	.0122***	.0125***
	(.0039)	(.0044)	(.0043)
Applied R&D intensity	.0054**	.0193***	.0196***
	(.0024)	(.0046)	(.0047)
Extramural R&D intensity	-.0009	.0015	.0015
	(.0016)	(.0009)	(.0009)
Capital R&D intensity	.0009**	.0045***	.0047***
	(.0004)	(.0010)	(.0010)
Number of control observations N_0	N_0= 5038	N_0= 7189	N_0 = 7195
Number of treated observations N_1	N_1 = 94784	N_1 = 101381	N_1 = 101641

Table A5. Sensitivity check: ATTs based on weights obtained through coarsed exact matching (CEM) by employment quartile.

	Growth of private R&D intensity	Growth of R&D personnel intensity	Growth of basic R&D intensity	Growth of experimental R&D intensity	Growth of applied R&D intensity	Growth of extramural R&D intensity	Growth of capital R&D intensity
Quartile 1	.0993***	.0681***	.0139**	.0389***	.0543***	.0047	.0116***
N_0= 35765	(.0233)	(.0146)	(.0054)	(.0091)	(.0145)	(.0029)	(.0032)
N_1 = 1081							
Quartile 2	.0043	.0142**	.0003	.0027*	.0041**	-.0035	.0008**
N_0= 33303	(.0033)	(.0072)	(.0002)	(.0015)	(.0020)	(.0036)	(.0003)
N_1 = 1781							
Quartile 3	.0179	.0022	.0021	.0161	-.0031	.0026	.0005
N_0= 33303	(.0158)	(.0263)	(.0022)	(.0607)	(.0125)	(.0020)	(.0004)
N_1 = 1781							
Quartile 4	.0011*	.0026	.0012	.0086	.0026	-.0003	.0016*
N_0= 30921	(.0006)	(.0021)	(.0036)	(.0244)	(.0021)	(.0004)	(.0009)
N_1 = 5028							

Note: N_0 - number of control observations, N_1 - number of treated observations.

Table A6. Regression-based indicators of effect-size heterogeneity

Dependent variable	Effect-size estimate based on growth of private R&D intensity	Effect-size estimate based on logarithm of private R&D intensity
Private R&D	0.0005***	0.0005***
	(0.0001)	(0.0001)
Private R&D squared	-0.0094***	-0.0094***
	(0.0012)	(0.0012)
Logarithm of firm age	-0.0016***	-0.0016***
	(0.0001)	(0.0001)
Logarithm of firm employment	-0.0003***	-0.0003***
	(0.0001)	(0.0001)
Interaction of logarithm of firm age*	-0.0002***	-0.0002***
logarithm of firm employment	(0.00003)	(0.00003)
Market share at 3-digit industry	-0.0151***	-0.0150***
SIC level	(0.0008)	(0.0008)
Distance to the R&D frontier	0.0005***	0.0005***
	(0.00004)	(0.00004)
Logarithm of firm turnover	-0.0002***	-0.0002***
	(0.00004)	(0.00004)
Logarithm of number of plants	-0.0002***	-0.0002***
	(0.00006)	(0.00006)
dotcom	-0.0010***	-0.0010***
	(0.0001)	(0.0001)
crisis2008_2010	-0.0011***	-0.0010***
	(0.00007)	(0.00007)
_cons	-0.0001	0.0681***
	(0.0004)	(0.0004)
N-observations	123152	123152
AIC	-751227.0	-751676.3
BIC	-751110.4	-751559.6
F-statistic	543.9	543.9

Note: The estimated treatment effect is regressed on firm characteristics and indicator variables for crisis periods. Heteroscedasticity and first order autocorrelation robust errors are in parentheses.
Dependent variable: Predicted ATT effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7. Entropy weighting balance for the entire sample in estimation of the growth effects of private R&D intensity.

covriates	mean_Tr	mean_Co_Pre	mean_Co_Post	var_Tr	var_Co_Pre	var_Co_Post	skew_Tr	skew_Co_Pre	skew_Co_Post	sdiff_Pre	sdiff_Post
Llogage	2.6587	2.8179	2.6588	0.5171	0.4329	0.5171	-0.8720	-1.0738	-0.9802	-0.2214	-0.0001
Llog_empl	3.1526	4.5059	3.1529	2.9029	4.2553	2.9040	0.4349	0.1909	0.5212	-0.7943	-0.0002
Luk_ownership	0.8697	0.7598	0.8697	0.1133	0.1825	0.1134	-2.1963	-1.2162	-2.1961	0.3265	0.0000
Lrd_int	0.0857	0.0706	0.0857	0.0210	0.0243	0.0210	3.0735	3.4375	3.3714	0.1044	0.0000
Lpri_rd_int	0.0768	0.0688	0.0768	0.0164	0.0232	0.0164	3.1403	3.4662	3.4598	0.0631	0.0000
Llive_lu	2.9742	11.1875	2.9763	3453.0116	18648.4516	3458.5298	84.7473	26.2460	79.3199	-0.1398	0.0000
Llogrturn	7.3712	9.0568	7.3717	5.3348	6.9857	5.3362	0.0474	0.0544	0.0597	-0.7298	-0.0002
Llogrprod	4.3725	4.6205	4.3727	0.9009	1.3776	0.9009	0.3437	-0.0608	0.1517	-0.2612	-0.0001
LRD_pers	6.0639	45.3304	6.0678	8453.6328	30356.3409	8453.3452	46.1794	9.0374	30.2408	-0.4271	0.0000
Lgrowth	0.0517	0.0764	0.0517	0.3038	0.3866	0.3038	1.2971	2.0933	0.8197	-0.0447	0.0000
Lshare	0.0064	0.0262	0.0064	0.0013	0.0074	0.0013	14.6850	6.6440	16.0548	-0.5433	-0.0001
LUK_sub	0.9674	0.4178	0.9674	0.0315	0.2433	0.0315	-5.2671	0.3333	-5.2670	3.0967	0.0000
Ldouble_ukeu	0.7787	0.2927	0.7787	0.1723	0.2071	0.1723	-1.3428	0.9110	-1.3430	1.1707	-0.0001
Lcivil_dummy	0.4278	0.7623	0.4278	0.2448	0.1812	0.2448	0.2920	-1.2323	0.2919	-0.6762	-0.0001
Lsmesrd	0.8493	0.5432	0.8492	0.1280	0.2482	0.1280	-1.9525	-0.1736	-1.9520	0.8554	0.0001
Llargerd	0.0670	0.2326	0.0671	0.0625	0.1785	0.0626	3.4632	1.2658	3.4617	-0.6623	-0.0002
Lstart	0.0508	0.0288	0.0508	0.0482	0.0279	0.0482	4.0899	5.6388	4.0901	0.1005	0.0000
Lyoung	0.1809	0.1250	0.1809	0.1482	0.1094	0.1482	1.6581	2.2683	1.6583	0.1453	0.0000
Lsmes	0.7126	0.5699	0.7125	0.2048	0.2451	0.2049	-0.9393	-0.2823	-0.9391	0.3152	0.0001
LExtramur_RD_int	0.0054	0.0063	0.0054	0.0004	0.0010	0.0004	15.1907	9.6290	18.6757	-0.0473	0.0000
Lcapexprndint	0.0050	0.0035	0.0050	0.0001	0.0002	0.0001	5.1663	8.9567	4.8095	0.1518	0.0000
Lsmesrd08	0.4134	0.1596	0.4134	0.2425	0.1341	0.2425	0.3518	1.8590	0.3519	0.5154	0.0000
Llargerd08	0.0125	0.0546	0.0125	0.0123	0.0517	0.0124	8.7769	3.9190	8.7698	-0.3794	-0.0002
Lsubsidysmes	0.8239	0.2897	0.8239	0.1451	0.2058	0.1451	-1.7009	0.9274	-1.7006	1.4027	0.0001
distance	0.0953	0.1238	0.0944	0.0109	0.0195	0.0108	3.1484	2.2940	3.2158	-0.2738	0.0081
herfindahl_index	0.0942	0.1230	0.0942	0.0108	0.0193	0.0108	3.1975	2.5043	3.1882	-0.2774	0.0000
crisis2008_2010	0	0	0	0	0	0	1	1	1	0	0
dotcom	0	0	0	0	0	0	2	2	2	0	0
pavittd1	0	0	0	0	0	0	1	2	1	0	0
pavittd2	0	0	0	0	0	0	1	3	1	0	0
pavittd3	0	0	0	0	0	0	3	2	3	0	0
pavittd4	0	0	0	0	0	0	1	0	1	0	0

2000.year	0	0	0	0	0	0	5	3	5	0	0
2001.year	0	0	0	0	0	0	5	2	5	0	0
2003.year	0	0	0	0	0	0	5	5	5	0	0
2004.year	0	0	0	0	0	0	4	6	4	0	0
2005.year	0	0	0	0	0	0	4	4	4	0	0
2006.year	0	0	0	0	0	0	3	3	3	0	0
2007.year	0	0	0	0	0	0	3	2	3	0	0
2008.year	0	0	0	0	0	0	3	1	3	0	0
2009.year	0	0	0	0	0	0	3	5	3	0	0
2011.year	0	0	0	0	0	0	3	4	3	0	0
2.sic07_2d	0	0	0	0	0	0	45	72	45	0	0
5.sic07_2d	0	0	0	0	0	0	106	42	106	0	0
6.sic07_2d	0	0	0	0	0	0	68	15	68	0	0
8.sic07_2d	0	0	0	0	0	0	26	15	26	0	0
9.sic07_2d	0	0	0	0	0	0	31	22	31	0	0
10.sic07_2d	0	0	0	0	0	0	5	7	5	0	0
11.sic07_2d	0	0	0	0	0	0	19	24	19	0	0
12.sic07_2d	0	0	0	0	0	0	102	21	102	0	0
13.sic07_2d	0	0	0	0	0	0	9	7	9	0	0
14.sic07_2d	0	0	0	0	0	0	16	11	16	0	0
15.sic07_2d	0	0	0	0	0	0	21	20	21	0	0
16.sic07_2d	0	0	0	0	0	0	15	10	15	0	0
17.sic07_2d	0	0	0	0	0	0	12	8	12	0	0
18.sic07_2d	0	0	0	0	0	0	11	6	11	0	0
19.sic07_2d	0	0	0	0	0	0	43	11	43	0	0
20.sic07_2d	0	0	0	0	0	0	5	6	5	0	0
21.sic07_2d	0	0	0	0	0	0	15	8	15	0	0
22.sic07_2d	0	0	0	0	0	0	6	4	6	0	0
23.sic07_2d	0	0	0	0	0	0	8	12	8	0	0
24.sic07_2d	0	0	0	0	0	0	15	6	15	0	0
25.sic07_2d	0	0	0	0	0	0	4	3	4	0	0
26.sic07_2d	0	0	0	0	0	0	3	5	3	0	0
27.sic07_2d	0	0	0	0	0	0	6	11	6	0	0
28.sic07_2d	0	0	0	0	0	0	4	6	4	0	0
29.sic07_2d	0	0	0	0	0	0	7	7	7	0	0

30.sic07_2d	0	0	0	0	0	0	13	7	13	0	0
31.sic07_2d	0	0	0	0	0	0	10	5	10	0	0
32.sic07_2d	0	0	0	0	0	0	6	4	6	0	0
33.sic07_2d	0	0	0	0	0	0	9	7	9	0	0
35.sic07_2d	0	0	0	0	0	0	49	17	49	0	0
36.sic07_2d	0	0	0	0	0	0	50	18	50	0	0
37.sic07_2d	0	0	0	0	0	0	49	22	49	0	0
38.sic07_2d	0	0	0	0	0	0	31	15	31	0	0
41.sic07_2d	0	0	0	0	0	0	13	31	13	0	0
42.sic07_2d	0	0	0	0	0	0	18	32	18	0	0
43.sic07_2d	0	0	0	0	0	0	9	18	9	0	0
45.sic07_2d	0	0	0	0	0	0	11	28	11	0	0
46.sic07_2d	0	0	0	0	0	0	3	4	3	0	0
47.sic07_2d	0	0	0	0	0	0	6	11	6	0	0
49.sic07_2d	0	0	0	0	0	0	34	4	34	-1	0
50.sic07_2d	0	0	0	0	0	0	110	12	110	-1	0
51.sic07_2d	0	0	0	0	0	0	122	24	122	0	0
52.sic07_2d	0	0	0	0	0	0	28	4	28	-1	0
53.sic07_2d	0	0	0	0	0	0	54	11	54	0	0
55.sic07_2d	0	0	0	0	0	0	16	42	16	0	0
56.sic07_2d	0	0	0	0	0	0	13	27	13	0	0
58.sic07_2d	0	0	0	0	0	0	7	7	7	0	0
59.sic07_2d	0	0	0	0	0	0	11	13	11	0	0
61.sic07_2d	0	0	0	0	0	0	16	13	16	0	0
62.sic07_2d	0	0	0	0	0	0	3	4	3	0	0
63.sic07_2d	0	0	0	0	0	0	18	20	18	0	0
64.sic07_2d	0	0	0	0	0	0	21	25	21	0	0
65.sic07_2d	0	0	0	0	0	0	27	31	27	0	0
66.sic07_2d	0	0	0	0	0	0	10	21	10	0	0
68.sic07_2d	0	0	0	0	0	0	11	17	11	0	0
69.sic07_2d	0	0	0	0	0	0	11	24	11	0	0
70.sic07_2d	0	0	0	0	0	0	6	16	6	0	0
71.sic07_2d	0	0	0	0	0	0	4	7	4	0	0
72.sic07_2d	0	0	0	0	0	0	2	4	2	0	0
73.sic07_2d	0	0	0	0	0	0	11	24	11	0	0

74.sic07_2d	0	0	0	0	0	0	12	36	12	0	0
75.sic07_2d	0	0	0	0	0	0	38	42	38	0	0
77.sic07_2d	0	0	0	0	0	0	12	34	12	0	0
78.sic07_2d	0	0	0	0	0	0	13	20	13	0	0
79.sic07_2d	0	0	0	0	0	0	18	6	18	0	0
81.sic07_2d	0	0	0	0	0	0	16	59	16	0	0
82.sic07_2d	0	0	0	0	0	0	7	16	7	0	0
84.sic07_2d	0	0	0	0	0	0	366	46	366	0	0
85.sic07_2d	0	0	0	0	0	0	16	5	16	-1	0
86.sic07_2d	0	0	0	0	0	0	18	14	18	0	0
87.sic07_2d	0	0	0	0	0	0	31	25	31	0	0
88.sic07_2d	0	0	0	0	0	0	30	14	30	0	0
90.sic07_2d	0	0	0	0	0	0	15	15	15	0	0
91.sic07_2d	0	0	0	0	0	0	35	14	35	0	0
92.sic07_2d	0	0	0	0	0	0	59	34	59	0	0
93.sic07_2d	0	0	0	0	0	0	23	13	23	0	0
94.sic07_2d	0	0	0	0	0	0	35	11	35	0	0
95.sic07_2d	0	0	0	0	0	0	21	27	21	0	0
96.sic07_2d	0	0	0	0	0	0	11	10	11	0	0
1.Lcivil_dummy#c.Lrd_int	0.0249	0.0629	0.0249	0.0072	0.0228	0.0072	6.2998	3.6120	7.0344	-0.4497	0.0000
1.Lyoung#c.Lrd_int	0.0279	0.0101	0.0279	0.0102	0.0043	0.0102	5.2063	9.6035	5.3884	0.1756	0.0000
1.Lsmes#c.Lrd_int	0.0484	0.0386	0.0484	0.0119	0.0164	0.0119	4.3930	4.7433	4.8778	0.0904	0.0000
1.Lcivil_dummy#c.Lpri_rd_int	0.0228	0.0617	0.0228	0.0059	0.0221	0.0059	6.3436	3.6310	6.9623	-0.5086	0.0000
1.Lyoung#c.Lpri_rd_int	0.0244	0.0095	0.0244	0.0079	0.0038	0.0079	5.3442	9.9210	5.7208	0.1682	0.0000
1.Lsmes#c.Lpri_rd_int	0.0440	0.0376	0.0440	0.0093	0.0158	0.0093	4.3628	4.7907	4.7373	0.0668	0.0000
1.Lcivil_dummy#c.LRD_pers	3.0453	43.3404	3.0489	3733.5010	29279.9351	3733.4975	68.5748	9.1740	50.7426	-0.6595	-0.0001
1.Lyoung#c.LRD_pers	0.2520	0.9973	0.2521	22.0997	125.1289	22.1041	247.7113	18.2221	79.0989	-0.1585	0.0000
1.Lsmes#c.LRD_pers	1.8768	6.1555	1.8769	28.6453	640.1904	28.6743	15.6480	6.7102	17.6625	-0.7994	0.0000

Notes: Number of treated N_1 = 133563 observations, and number of control observations N_0 = 10282. Industry and year dummies are excluded to save space. Mean_Tr: mean value of the covariate for treated firms; Mean_Pre: mean value of the covariate in the control group before balancing; Mean_Post: mean value of the covariate in the control group after balancing; Var_Tr: variance of the covariate for treated group; Var_Pre: variance of the covariate for the control group before balancing; Var_Post: variance of the covariate for the control group after balancing; Skew_Tr: skewness of the covariate for the treated group; Skew_Pre: skewness of the covariate for the control group before balancing; Skew_Post: skewness of the covariate for the control group after balancing; SD_diff_Pre: standardized difference between treated and control group of firms before entropy balancing; SD_diff_Post: standardized difference after entropy balancing. The standardized difference in means between the treatment and control group is given by $(\bar{s}_1 - \bar{s}_0)/\sqrt{1/2(\sigma_{s1}^2 + (\sigma_{s0}^2)}$, where $\bar{s}_1$ and $\bar{s}_0$ are covariate means in the treated and control groups; and σ_{s1}^2 and σ_{s0}^2 are corresponding covariate variances. For other notes, see Table A1.

Table A8. Entropy balance of the third quarter of the firms' employment distribution in estimation of the growth effects of private R&D intensity (industry dummies are omitted to save space).

covariates	mean_Tr	mean_Co_Pre	mean_Co_Post	var_Tr	var_Co_Pre	var_Co_Post	skew_Tr	skew_Co_Pre	skew_Co_Post	sdiff_Pre	sdiff_Post
Llogage	2.9119	2.9079	2.9121	0.3015	0.3369	0.3015	-1.0321	-1.1196	-1.1671	0.0073	-0.0003
Llog_empl	3.7879	3.8326	3.7881	0.1785	0.1739	0.1785	0.0749	-0.0189	-0.0497	-0.1059	-0.0006
Luk_ownership_dummy	0.8561	0.8490	0.8561	0.1232	0.1282	0.1232	-2.0288	-1.9499	-2.0294	0.0200	-0.0002
Lrd_int	0.0505	0.0361	0.0505	0.0078	0.0103	0.0078	4.7952	5.4015	2.9855	0.1633	-0.0001
Lpri_rd_int	1.3555	1.3503	1.3555	1.2831	0.9157	1.2832	6.8648	4.2791	5.2473	0.0046	-0.0001
Llive_lu	8.3039	8.3566	8.3045	0.8257	0.8966	0.8258	0.5168	0.4104	0.3769	-0.0579	-0.0006
Llogrturn	4.5412	4.5479	4.5415	0.6658	0.6911	0.6659	0.6675	0.6403	0.3100	-0.0082	-0.0004
Llogrprod	2.7989	2.5881	2.7993	12.7003	48.0609	12.7083	3.5471	5.5007	4.1885	0.0592	-0.0001
LRD_pers	0.0457	0.0658	0.0457	0.1817	0.2336	0.1817	1.5803	2.4123	0.0147	-0.0473	0.0000
Lgrowth	0.0034	0.0051	0.0034	0.0004	0.0006	0.0004	25.9244	23.1681	34.4906	-0.0880	0.0000
Lshare	0.9658	0.6201	0.9659	0.0330	0.2357	0.0330	-5.1294	-0.4950	-5.1303	1.9033	-0.0001
LUK_sub	0.7622	0.4588	0.7623	0.1812	0.2484	0.1813	-1.2320	0.1653	-1.2324	0.7128	-0.0002
Ldouble_ukeu	0.4410	0.6961	0.4411	0.2465	0.2116	0.2466	0.2375	-0.8527	0.2374	-0.5137	0.0000
Lcivil_dummy	0.9466	0.8313	0.9466	0.0505	0.1403	0.0506	-3.9728	-1.7691	-3.9738	0.5129	-0.0001
Lsmesrd	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	45.5981	44.9889	45.6024	-0.0006	0.0000
Llargerd	0.0106	0.0133	0.0106	0.0105	0.0131	0.0105	9.5737	8.4904	9.5744	-0.0269	0.0000
Lstart	0.0704	0.0849	0.0704	0.0655	0.0777	0.0655	3.3572	2.9795	3.3575	-0.0563	0.0000
Lyoung	0.0032	0.0032	0.0032	0.0002	0.0005	0.0002	17.1032	13.6465	16.4893	-0.0008	0.0000
LExtramur_RD_int	0.0031	0.0017	0.0031	0.0000	0.0001	0.0000	6.7348	16.5345	4.1781	0.2250	-0.0001
Lcapexprndint	0.4691	0.2131	0.4691	0.2491	0.1678	0.2492	0.1240	1.4011	0.1239	0.5129	0.0000
Lsmesrd08	0.9144	0.5525	0.9145	0.0783	0.2474	0.0783	-2.9629	-0.2113	-2.9637	1.2936	-0.0001
Lsubsidysmes	0.2705	0.1180	0.2705	0.0629	0.0334	0.0629	0.8886	2.6080	0.8814	0.6079	-0.0001
distance	0.0913	0.1199	0.0913	0.0108	0.0179	0.0108	3.2147	2.2149	3.4862	-0.2750	-0.0001
herfindahl_index	0.3126	0.4406	0.3126	0.2149	0.2466	0.2150	0.8084	0.2395	0.8085	-0.2760	0.0000
crisis2008_2010	0	0	0	0	0	0	3	2	3	0	0
dotcom	0	0	0	0	0	0	1	3	1	0	0
pavittd1	0	0	0	0	0	0	1	3	1	0	0
pavittd2	0	0	0	0	0	0	2	2	2	0	0
pavittd3	0	1	0	0	0	0	1	0	1	-1	0
pavittd4	0	0	0	0	0	0	6	3	6	0	0
2000.year	0	0	0	0	0	0	5	7	5	0	0
2001.year	0	0	0	0	0	0	4	12	4	0	0
2003.year	0	0	0	0	0	0	4	4	4	0	0

2004.year	0	0	0	0	0	0	3	4	3	0	0
2005.year	0	0	0	0	0	0	3	2	3	0	0
2006.year	0	0	0	0	0	0	3	6	3	0	0
2007.year	0	0	0	0	0	0	2	3	2	0	0
2009.year	0	0	0	0	0	0	2	3	2	0	0
2011.year	0	0	0	0	0	0	8	6	5	0	0
1.Lcivil_dummy#c.Lrd_int	0.0062	0.0042	0.0062	0.0019	0.0013	0.0019	11.5878	14.6160	8.9266	0.0457	0.0000
1.Lyoung#c.Lrd_int	0.0181	0.0277	0.0181	0.0037	0.0090	0.0037	7.7271	6.0251	5.2788	-0.1585	-0.0001
1.Lcivil_dummy#c.Lpri_rd_int	0.0058	0.0041	0.0058	0.0016	0.0013	0.0016	11.6997	15.0616	9.1517	0.0419	0.0000
1.Lyoung#c.Lpri_rd_int	0.0477	0.0349	0.0477	0.0067	0.0098	0.0067	4.7433	5.5215	2.9764	0.1560	-0.0001
1.Lcivil_dummy#c.LRD_pers	1.0498	2.1547	1.0501	7.9501	47.6062	7.9585	6.3161	5.7088	8.0313	-0.3919	-0.0001
1.Lyoung#c.LRD_pers	0.2371	0.1908	0.2371	2.2599	2.3991	2.2600	13.2226	20.9552	16.2129	0.0308	0.0000

Notes: Number of treated and control observations are 33562 and 2381 respectively. Industry and year dummies are excluded to save space. Mean_Tr: mean value of the covariate for treated firms; Mean_Pre: mean value of the covariate in the control group before balancing; Mean_Post: mean value of the covariate in the control group after balancing; Var_Tr: variance of the covariate for treated group; Var_Pre: variance of the covariate for the control group before balancing; Var_Post: variance of the covariate for the control group after balancing; Skew_Tr: skewness of the covariate for the treated group; Skew_Pre: skewness of the covariate for the control group before balancing; Skew_Post: skewness of the covariate for the control group after balancing; SD_diff_Pre: standardized difference between treated and control group of firms before entropy balancing; SD_diff_Post: standardized difference after entropy balancing. The standardized difference in means between the treatment and control group is given by $(\bar{s}_1 - \bar{s}_0)/\sqrt{1/2(\sigma_{s1}^2 + (\sigma_{s0}^2)}$, where $\bar{s}_1$ and $\bar{s}_0$ are covariate means in the treated and control groups; and σ_{s1}^2 and σ_{s0}^2 are corresponding covariate variances. For other notes, see Table A1.

Table A9. Entropy balance for the third quarter of the firms' deflated turnover distribution in estimation of the growth effects of private R&D intensity (industry dummies are omitted to save space).

covariates	mean_Tr	mean_Co_Pre	mean_Co_Post	var_Tr	var_Co_Pre	var_Co_Post	skew_Tr	skew_Co_Pre	skew_Co_Post	sdiff_Pre	sdiff_Post
Llogage	2.9427	2.9127	2.9429	0.2861	0.3002	0.2862	-1.0678	-1.1260	-0.9939	0.0561	-0.0002
Llog_empl	4.1609	4.3226	4.1611	0.6980	0.6523	0.6980	-0.1680	-0.1873	-0.0916	-0.1935	-0.0002
Luk_ownership_dummy	0.7991	0.7695	0.7992	0.1605	0.1774	0.1606	-1.4931	-1.2800	-1.4934	0.0739	-0.0001
Lrd_int	0.0344	0.0676	0.0344	0.0040	0.0265	0.0040	5.9130	3.4406	5.5202	-0.5279	0.0000
Lpri_rd_int	0.0326	0.0664	0.0326	0.0035	0.0259	0.0035	5.9223	3.4626	5.4499	-0.5707	0.0000
Llive_lu	1.7408	1.7100	1.7409	8.5906	6.7158	8.5909	31.1402	19.8674	17.5274	0.0105	0.0000
Llogrturn	8.8788	9.0372	8.8792	0.8723	0.9847	0.8724	0.3464	0.3323	0.7253	-0.1697	-0.0004
Llogrprod	4.7411	4.7339	4.7413	0.5539	0.5646	0.5540	0.6551	0.8015	0.9058	0.0097	-0.0003
LRD_pers	3.7659	12.3277	3.7661	49.4940	1001.8238	49.4991	10.1009	4.4402	17.5068	-1.2170	0.0000
Lgrowth	0.0589	0.0864	0.0589	0.1770	0.2410	0.1770	2.6882	3.3405	2.6598	-0.0655	0.0000
Lshare	0.0055	0.0075	0.0055	0.0006	0.0007	0.0006	21.2166	15.0327	27.0451	-0.0835	0.0000
LUK_sub	0.9651	0.5640	0.9651	0.0337	0.2460	0.0337	-5.0692	-0.2581	-5.0698	2.1858	0.0000
Ldouble_ukeu	0.7603	0.4070	0.7604	0.1822	0.2414	0.1823	-1.2197	0.3787	-1.2199	0.8278	-0.0001
Lcivil_dummy	0.4837	0.7585	0.4837	0.2497	0.1832	0.2498	0.0653	-1.2079	0.0652	-0.5499	0.0000
Lsmesrd	0.8988	0.7350	0.8988	0.0910	0.1948	0.0910	-2.6444	-1.0648	-2.6449	0.5431	-0.0001

Llargerd	0.0381	0.0476	0.0381	0.0367	0.0454	0.0367	4.8235	4.2492	4.8237	-0.0495	0.0000
Lstart	0.0085	0.0113	0.0085	0.0085	0.0112	0.0085	10.6848	9.2382	10.6853	-0.0303	0.0000
Lyoung	0.0618	0.0723	0.0618	0.0580	0.0671	0.0580	3.6407	3.3035	3.6408	-0.0436	0.0000
LExtramur_RD_int	0.0022	0.0050	0.0022	0.0001	0.0009	0.0001	23.5126	10.2385	23.9285	-0.2487	0.0000
Lcapexprndint	0.0021	0.0028	0.0021	0.0000	0.0001	0.0000	9.6762	9.9150	5.2973	-0.1738	-0.0001
Lsmesrd08	0.4756	0.2456	0.4757	0.2494	0.1853	0.2495	0.0976	1.1822	0.0975	0.4607	0.0000
Llargerd08	0.8675	0.4586	0.8675	0.1150	0.2484	0.1150	-2.1673	0.1660	-2.1678	1.2056	-0.0001
Lsubsidysmes	0.2713	0.1408	0.2713	0.0631	0.0387	0.0631	0.9318	2.0875	0.8524	0.5198	0.0000
distance	0.0917	0.1119	0.0917	0.0105	0.0160	0.0105	3.1936	2.3634	3.6199	-0.1983	0.0000
herfindahl_index	0.3098	0.4038	0.3098	0.2138	0.2408	0.2139	0.8226	0.3922	0.8226	-0.2032	0.0000
crisis2008_2010	0	0	0	0	0	0	3	2	3	0	0
dotcom	0	0	0	0	0	0	1	2	1	0	0
pavittd1	0	0	0	0	0	0	1	3	1	0	0
pavittd2	0	0	0	0	0	0	2	2	2	0	0
pavittd3	0	1	0	0	0	0	0	0	0	0	0
pavittd4	0	0	0	0	0	0	6	2	6	-1	0
2000.year	0	0	0	0	0	0	5	7	5	0	0
2001.year	0	0	0	0	0	0	4	8	4	0	0
2003.year	0	0	0	0	0	0	4	4	4	0	0
2004.year	0	0	0	0	0	0	3	5	3	0	0
2005.year	0	0	0	0	0	0	3	3	3	0	0
2006.year	0	0	0	0	0	0	3	4	3	0	0
2007.year	0	0	0	0	0	0	2	3	2	0	0
2009.year	0	0	0	0	0	0	2	3	2	0	0
2011.year	0	0	0	0	0	0	2	4	2	0	0
1.Lcivil_dummy#c.Lrd_int	0.0150	0.0606	0.0151	0.0025	0.0254	0.0025	8.5316	3.6063	9.0966	-0.9065	0.0000
1.Lyoung#c.Lrd_int	0.0032	0.0052	0.0032	0.0007	0.0026	0.0007	16.4286	13.5736	11.8486	-0.0736	0.0000
1.Lcivil_dummy#c.Lpri_rd_int	0.0144	0.0597	0.0144	0.0023	0.0249	0.0023	8.6121	3.6247	8.8823	-0.9422	0.0000
1.Lyoung#c.Lpri_rd_int	0.0030	0.0051	0.0030	0.0007	0.0026	0.0007	16.7264	13.7661	11.9773	-0.0824	0.0000
1.Lcivil_dummy#c.LRD_pers	1.8081	11.2315	1.8082	36.8520	970.5225	36.8559	12.7770	4.6411	26.7112	-1.5523	0.0000
1.Lyoung#c.LRD_pers	0.2482	0.6322	0.2482	4.6564	43.9295	4.6576	22.9246	15.0562	24.0826	-0.1780	0.0000

Notes: Number of treated and control observations are 33649 and N0= 2390 respectively. Industry and year dummies are excluded to save space. Mean_Tr: mean value of the covariate for treated firms; Mean_Pre: mean value of the covariate in the control group before balancing; Mean_Post: mean value of the covariate in the control group after balancing; Var_Tr: variance of the covariate for treated group; Var_Pre: variance of the covariate for the control group before balancing; Var_Post: variance of the covariate for the control group after balancing; Skew_Tr: skewness of the covariate for the treated group; Skew_Pre: skewness of the covariate for the control group before balancing; Skew_Post: skewness of the covariate for the control group after balancing; SD_diff_Pre: standardized difference between treated and control group of firms before entropy balancing; SD_diff_Post: standardized difference after entropy balancing. The standardized difference in means between the treatment and control group is given by $(\bar{s}_1 - \bar{s}_0)/\sqrt{1/2(\sigma^2_{s1} + (\sigma^2_{s0})}$, where $\bar{s}_1$ and $\bar{s}_0$ are covariate means in the treated and control groups; and σ^2_{s1} and σ^2_{s0} are corresponding covariate variances. For other notes, see Table A1.

Table A10. Entropy balance for the fourth quarter of the firms' market share distribution in estimation of the growth effects of private R&D intensity (industry dummies are omitted to save space).

covariates	mean_Tr	mean_Co_Pre	mean_Co_Post	var_Tr	var_Co_Pre	var_Co_Post	skew_Tr	skew_Co_Pre	skew_Co_Post	sdiff_Pre	sdiff_Post
Llogage	2.9830	2.9971	2.9831	0.2915	0.2822	0.2915	-1.2019	-1.2609	-1.1982	-0.0260	-0.0002
Llog_empl	4.9274	5.8180	4.9276	1.9534	2.7985	1.9535	0.2450	-0.0150	0.2811	-0.6372	-0.0002
Luk_ownership_dummy	0.7145	0.6267	0.7146	0.2040	0.2340	0.2040	-0.9500	-0.5237	-0.9502	0.1946	-0.0001
Lrd_int	0.0283	0.0730	0.0284	0.0045	0.0227	0.0045	6.9835	3.4634	6.6133	-0.6642	-0.0001
Lpri_rd_int	0.0258	0.0721	0.0258	0.0035	0.0221	0.0035	7.1696	3.4701	6.4556	-0.7870	-0.0001
Llive_lu	8.6336	20.4338	8.6372	13246.8217	35899.9787	13254.5278	43.2826	18.8667	34.8862	-0.1025	0.0000
Llogrturn	9.7360	10.7821	9.7364	2.6902	4.2993	2.6904	0.2483	0.0511	0.1519	-0.6378	-0.0002
Llogrprod	4.8291	4.9773	4.8293	0.8378	1.2538	0.8378	1.2518	0.3458	0.8905	-0.1619	-0.0002
LRD_pers	18.4286	82.0478	18.4342	32289.7411	55326.5154	32302.8325	23.6135	6.6356	15.0057	-0.3540	0.0000
Lgrowth	0.0677	0.1098	0.0677	0.2289	0.3539	0.2289	3.9279	4.6761	4.5892	-0.0879	0.0000
Lshare	0.0231	0.0498	0.0231	0.0046	0.0131	0.0046	7.6969	4.8089	9.7006	-0.3928	0.0000
LUK_sub	0.9477	0.3412	0.9477	0.0496	0.2248	0.0496	-4.0202	0.6701	-4.0203	2.7233	0.0000
Ldouble_ukeu	0.7323	0.2531	0.7323	0.1961	0.1891	0.1961	-1.0491	1.1356	-1.0493	1.0821	-0.0001
Lcivil_dummy	0.5973	0.8299	0.5973	0.2405	0.1412	0.2406	-0.3968	-1.7560	-0.3970	-0.4742	-0.0001
Lsmesrd	0.6334	0.3333	0.6334	0.2322	0.2223	0.2322	-0.5537	0.7071	-0.5538	0.6227	0.0000
Llargerd	0.2354	0.4156	0.2354	0.1800	0.2429	0.1800	1.2474	0.3425	1.2474	-0.4247	0.0000
Lstart	0.0094	0.0091	0.0094	0.0093	0.0090	0.0093	10.1938	10.3277	10.1941	0.0024	0.0000
Lyoung	0.0581	0.0545	0.0581	0.0547	0.0516	0.0547	3.7778	3.9237	3.7779	0.0153	0.0000
Lsmes	0.7254	0.4413	0.7254	0.1992	0.2466	0.1992	-1.0099	0.2365	-1.0101	0.6365	-0.0001
LExtramur_RD_int	0.0026	0.0087	0.0026	0.0003	0.0015	0.0003	22.2118	7.9446	19.6803	-0.3786	-0.0001
Lcapexprndint	0.0017	0.0040	0.0017	0.0000	0.0002	0.0000	14.6299	8.6990	9.3137	-0.4394	-0.0001
Lsmesrd08	0.3656	0.1199	0.3656	0.2319	0.1055	0.2320	0.5583	2.3408	0.5583	0.5102	0.0000
Llargerd08	0.0475	0.1044	0.0475	0.0453	0.0935	0.0453	4.2524	2.5873	4.2524	-0.2672	0.0000
Lsubsidysmes	0.6017	0.1824	0.6017	0.2397	0.1492	0.2397	-0.4154	1.6449	-0.4155	0.8564	0.0000
mrkt_share_3d	0.0238	0.0527	0.0238	0.0046	0.0146	0.0046	7.6966	4.6286	9.2873	-0.4254	0.0000
herfindahl_index	0.1068	0.1294	0.1068	0.0127	0.0204	0.0127	2.9572	2.5824	3.1867	-0.2004	0.0000
crisis2008_2010	0	0	0	0	0	0	1	1	1	0	0
dotcom	0	0	0	0	0	0	2	2	2	0	0
pavittd1	0	0	0	0	0	0	2	2	2	0	0
pavittd2	0	0	0	0	0	0	1	2	1	0	0
pavittd3	0	0	0	0	0	0	2	2	2	0	0
pavittd4	0	0	0	0	0	0	0	1	0	0	0
2000.year	0	0	0	0	0	0	5	3	5	0	0

2001.year	0	0	0	0	0	0	4	4	4	0	0
2003.year	0	0	0	0	0	0	4	5	4	0	0
2004.year	0	0	0	0	0	0	4	5	4	0	0
2005.year	0	0	0	0	0	0	4	4	4	0	0
2006.year	0	0	0	0	0	0	3	3	3	0	0
2007.year	0	0	0	0	0	0	3	2	3	0	0
2008.year	0	0	0	0	0	0	3	4	3	0	0
2009.year	0	0	0	0	0	0	3	4	3	0	0
2011.year	0	0	0	0	0	0	2	4	2	0	0
1.Lcivil_dummy#c.Lrd_int	0.0148	0.0686	0.0148	0.0027	0.0221	0.0027	9.5565	3.5685	10.6283	-1.0341	-0.0001
1.Lyoung#c.Lrd_int	0.0033	0.0040	0.0033	0.0010	0.0015	0.0010	17.1006	15.1994	14.7812	-0.0227	0.0000
1.Lsmes#c.Lrd_int	0.0204	0.0284	0.0204	0.0031	0.0118	0.0031	8.3263	5.6892	7.9586	-0.1457	0.0000
1.Lcivil_dummy#c.Lpri_rd_int	0.0139	0.0679	0.0139	0.0023	0.0216	0.0023	9.6356	3.5628	10.1075	-1.1303	-0.0001
1.Lyoung#c.Lpri_rd_int	0.0029	0.0038	0.0029	0.0008	0.0014	0.0008	17.5663	15.6377	15.8901	-0.0326	0.0000
1.Lsmes#c.Lpri_rd_int	0.0189	0.0279	0.0189	0.0025	0.0115	0.0025	8.3588	5.7200	7.5945	-0.1796	0.0000
1.Lcivil_dummy#c.LRD_pers	9.6741	78.5413	9.6795	14278.6253	53474.6095	14291.2895	35.1377	6.7318	25.6469	-0.5763	0.0000
1.Lyoung#c.LRD_pers	0.3780	1.5699	0.3781	79.9229	219.8841	79.9274	138.4234	14.3734	46.1194	-0.1333	0.0000
1.Lsmes#c.LRD_pers	2.8497	7.4422	2.8499	57.4073	845.1896	57.4696	13.2679	6.0998	15.1341	-0.6061	0.0000

Notes: Number of treated and control observations are 30396 and 5006 respectively. Industry and year dummies are excluded to save space. Mean_Tr: mean value of the covariate for treated firms; Mean_Pre: mean value of the covariate in the control group before balancing; Mean_Post: mean value of the covariate in the control group after balancing; Var_Tr: variance of the covariate for treated group; Var_Pre: variance of the covariate for the control group before balancing; Var_Post: variance of the covariate for the control group after balancing; Skew_Tr: skewness of the covariate for the treated group; Skew_Pre: skewness of the covariate for the control group before balancing; Skew_Post: skewness of the covariate for the control group after balancing; SD_diff_Pre: standardized difference between treated and control group of firms before entropy balancing; SD_diff_Post: standardized difference after entropy balancing. The standardized difference in means between the treatment and control group is given by $(\bar{s}_1 - \bar{s}_0)/\sqrt{1/2(\sigma_{s1}^2 + (\sigma_{s0}^2)}$, where $\bar{s}_1$ and $\bar{s}_0$ are covariate means in the treated and control groups; and σ_{s1}^2 and σ_{s0}^2 are corresponding covariate variances. For other notes, see Table A1.

Table A11. Entropy balance for the first quarter of distance to R&D intensity frontier in estimation of the growth effects of private R&D intensity (industry dummies are omitted to save space).

covariates	mean_Tr	mean_Co_Pre	mean_Co_Post	var_Tr	var_Co_Pre	var_Co_Post	skew_Tr	skew_Co_Pre	skew_Co_Post	sdiff_Pre	sdiff_Post
Llogage	2.8668	2.9062	2.8668	0.3723	0.3613	0.3724	-1.1068	-1.2144	-0.9547	-0.0646	-0.0001
Llog_empl	3.7027	3.8486	3.7029	1.2237	0.9798	1.2243	-0.4554	-0.1830	-0.3955	-0.1319	-0.0002
Luk_ownership	0.8606	0.8411	0.8606	0.1200	0.1337	0.1200	-2.0816	-1.8663	-2.0818	0.0561	0.0000
Lrd_int	0.0406	0.0389	0.0407	0.0108	0.0153	0.0109	5.4991	5.1682	5.1354	0.0171	-0.0004
Lpri_rd_int	0.0378	0.0380	0.0378	0.0091	0.0149	0.0092	5.5060	5.2233	5.3098	-0.0018	-0.0004
Llive_lu	1.5489	1.5884	1.5490	5.1089	7.7647	5.1138	17.9752	19.7471	7.7937	-0.0175	-0.0001
Llogrturn	8.0502	8.3788	8.0507	2.8061	2.3115	2.8072	-0.6772	0.1190	-0.6663	-0.1962	-0.0003
Llogrprod	4.4041	4.5690	4.4044	0.8905	0.8920	0.8906	0.0504	0.8342	-0.1817	-0.1747	-0.0003
LRD_pers	1.7415	6.3076	1.7441	25.9177	601.1121	26.0527	25.4581	7.0233	15.1790	-0.8969	-0.0005

Lgrowth	0.0297	0.0598	0.0297	0.2179	0.2505	0.2179	0.5194	2.9018	-1.6035	-0.0644	0.0000
Lshare	0.0073	0.0090	0.0073	0.0012	0.0014	0.0012	16.3841	12.0098	17.2011	-0.0516	0.0000
LUK_sub	0.9426	0.6219	0.9426	0.0541	0.2352	0.0541	-3.8069	-0.5026	-3.8059	1.3794	0.0001
Ldouble_ukeu	0.6584	0.4942	0.6584	0.2249	0.2501	0.2250	-0.6681	0.0230	-0.6679	0.3462	0.0001
Lcivil_dummy	0.6181	0.7559	0.6183	0.2361	0.1846	0.2361	-0.4863	-1.1917	-0.4869	-0.2836	-0.0003
Lsmesrd	0.9040	0.8221	0.9041	0.0868	0.1463	0.0868	-2.7433	-1.6842	-2.7438	0.2782	-0.0001
Llargerd	0.0202	0.0277	0.0203	0.0198	0.0269	0.0199	6.8120	5.7583	6.8105	-0.0527	-0.0001
Lstart	0.0184	0.0183	0.0184	0.0180	0.0180	0.0180	7.1762	7.1811	7.1737	0.0002	-0.0001
Lyoung	0.1016	0.0938	0.1016	0.0913	0.0850	0.0913	2.6372	2.7861	2.6368	0.0258	-0.0001
LExtramur_RD_int	0.0033	0.0032	0.0033	0.0003	0.0005	0.0003	16.2328	13.6809	17.0108	0.0082	-0.0004
Lcapexprndint	0.0026	0.0020	0.0026	0.0001	0.0001	0.0001	8.4615	13.1888	10.1238	0.0711	-0.0001
Lsmesrd08	0.3674	0.2455	0.3674	0.2324	0.1853	0.2325	0.5502	1.1826	0.5499	0.2528	-0.0001
Llargerd08	0.0009	0.0022	0.0009	0.0009	0.0022	0.0009	32.4478	21.4632	32.4556	-0.0393	0.0000
Lsubsidysmes	0.8523	0.5557	0.8524	0.1259	0.2470	0.1259	-1.9864	-0.2243	-1.9865	0.8361	0.0000
distance	0.0300	0.0268	0.0300	0.0004	0.0005	0.0004	0.0938	0.3760	0.1773	0.1540	0.0001
herfindahl_index	0.1224	0.1455	0.1224	0.0161	0.0226	0.0161	2.6662	2.0550	3.2025	-0.1820	0.0000
crisis2008_2010	0	1	0	0	0	0	1	0	1	-1	0
dotcom	0	0	0	0	0	0	2	2	2	0	0
pavittd1	0	0	0	0	0	0	3	6	3	0	0
pavittd2	0	0	0	0	0	0	2	4	2	0	0
pavittd3	0	0	0	0	0	0	2	2	2	0	0
pavittd4	1	1	1	0	0	0	0	-1	0	0	0
2000.year	0	0	0	0	0	0	5	3	5	0	0
2001.year	0	0	0	0	0	0	4	7	4	0	0
2003.year	0	0	0	0	0	0	4	9	4	0	0
2004.year	0	0	0	0	0	0	4	5	4	0	0
2005.year	0	0	0	0	0	0	3	4	3	0	0
2006.year	0	0	0	0	0	0	2	3	2	0	0
2007.year	0	0	0	0	0	0	3	1	3	-1	0
2009.year	0	0	0	0	0	0	3	5	3	0	0
2011.year	0	0	0	0	0	0	3	4	3	0	0

Notes: Number of treated and control observations are 22173 and 2782 respectively. Industry and year dummies are excluded to save space. Mean_Tr: mean value of the covariate for treated firms; Mean_Pre: mean value of the covariate in the control group before balancing; Mean_Post: mean value of the covariate in the control group after balancing; Var_Tr: variance of the covariate for treated group; Var_Pre: variance of the covariate for the control group before balancing; Var_Post: variance of the covariate for the control group after balancing; Skew_Tr: skewness of the covariate for the treated group; Skew_Pre: skewness of the covariate for the control group before balancing; Skew_Post: skewness of the covariate for the control group after balancing; SD_diff_Pre: standardized difference between treated and control group of firms before entropy balancing; SD_diff_Post: standardized difference after entropy balancing. The standardized difference in means between the treatment and control group is given by $(\bar{s}_1 - \bar{s}_0)/\sqrt{1/2(\sigma_{s1}^2 + (\sigma_{s0}^2)}$, where $\bar{s}_1$ and $\bar{s}_0$ are covariate means in the treated and control groups; and σ_{s1}^2 and σ_{s0}^2 are corresponding covariate variances. For other notes, see Table A1.