APPENDICES: FOR ONLINE PUBLICATION ONLY

Appendix A: Institutional details of policy and tax-adjusted user cost

A.1 SME definition

The UK Tax Relief Scheme's SME (Small and Medium Sized Enterprise) definition is set according to the European Commission's Recommendations 1996/280/EC (effective up to December 31st 2004) and 2003/361/EC (effective from January 1st 2005). The definition is based on assets ("balance sheet total"), employment ("staff headcount") and sales ("turnover") as described in Section 2. There are also some restrictions on the type of eligible enterprises. Below is a summary of the SME definition according to Recommendations 2003/361/EC. For further technical details on many parts of the tax rules see http://www.hmrc.gov.uk/manuals/cirdmanual/CIRD91400.htm.

Measurements of staff headcount, assets, and sales turnover for ceiling tests: The assets is the gross amount of assets shown in the accounts. The staff headcount of an enterprise represents the number of full-time person-years attributable to people who have worked within or for the enterprise during the year under consideration.[1] For the purposes of the sales turnover test, VAT and other indirect taxes are excluded from the figures. The staff headcount and financial data used for the "ceiling tests" (the maximum values possible for a firm to be eligible for SME status) are those relating to the latest approved accounting period, calculated on an annual basis and aggregated according to the rules described below. Assets and sales figures are converted to euros using the exchange rate on the last day of the relevant accounting period, or the average exchange rate throughout that accounting period (whichever is more beneficial for the enterprise).

An enterprise passes the ceiling tests if its staff headcount and either its aggregated assets or its aggregated turnover fall below the respective ceilings. An enterprise loses (acquires) its SME status if it fails (passes) the ceiling tests over two consecutive accounting periods.

Account aggregation rules for different enterprise types: In the case of an autonomous enterprise, the staff headcount and financial data are determined exclusively on the basis of the accounts of the enterprise (or the consolidated accounts) of the enterprise itself.[2] In the case of a linked enterprise, the ceiling tests are applied to the aggregates of the figures in its own accounts and those from the accounts of all other enterprises to which it is linked (including non-UK ones), unless the linked enterprises'

[1] The contributions of part-time workers, or those who work on a seasonal or temporary basis count as appropriate fractions of a full-time person-year. The term staff includes employees, persons seconded to the enterprise, owner-managers, partners (other than sleeping partners); it excludes apprentices or students engaged in vocational training with an apprenticeship or vocational training contract, and any periods or maternity or parental leave.

[2] An autonomous enterprise is defined by exclusion: one that is not a linked enterprise or a partner enterprise. Generally, an enterprise is autonomous if it has holding of less than 25% of the capital or voting rights in one or more enterprises and/or other enterprises do not have a stake of 25% or more of the capital voting rights in the enterprise.

account data are already included through account consolidation.[3] For further details of the European Commission's recommendations for SME definition, see Recommendations 1996/280/EC and 2003/361/EC.

A.2 UK R&D Tax Relief Scheme

The UK R&D Tax Relief Scheme includes a SME Scheme and a Large Company ("LCO") Scheme.[4] Since its introduction in 2000 up until 2012, more than 28,500 different companies had made claims under the SME Scheme, and over 7,000 under the Large Company Scheme, claiming more than £9.5bn in total R&D support. The annual amount of R&D support had risen to over £1bn by 2008, reaching £1.4bn in 2012, and covering qualifying R&D expenditure worth £13.2bn (HMRC Research and Development Tax Credit Statistics, 2014).

Both SME and Large Company Schemes are volume-based, i.e. the tax relief accrues on the total R&D spending rather than the incremental R&D over a prior base (the main US tax credit scheme is incremental). It works mostly through enhanced deduction of current R&D expenditure from taxable income, thus reducing R&D-performing companies' corporate tax liabilities.[5] In addition, under the SME Scheme, a company that has taxable loss after the additional deduction can also claim payable tax credit up to the amount of payable credit rate × enhanced qualifying R&D expenditure.[6] This payable tax credit can only be used to reduce the company's PAYE (Pay-As-You-Earn) or NIC (National Insurance Contributions) liabilities. Alternatively, the company (either as an SME or as a large company) can choose to carry the loss forward as normal.[7]

Qualifying R&D expenditure must be allowable as a deduction in calculating trading profits, which includes all flow costs, employee costs, staff providers, materials, payments to clinical trials volunteers, utilities, software, or subcontracted R&D expenditure (only if the contractor is an SME).[8] To be eligible for R&D tax relief, a company must also spend at least £10,000 a year on qualifying R&D

[3] Linked enterprises are those in which one enterprise is able to exercise control, directly or indirectly, over the affairs of the other.

[4] For further details, see http://www.hmrc.gov.uk/manuals/cirdmanual/CIRD90000.htm (SME Scheme) and http://www.hmrc.gov.uk/manuals/cirdmanual/CIRD85050.htm (Large Company Scheme).

[5] For example, if a company is allowed an enhancement rate of 75%, for each £100 of qualifying R&D expenditure it spends, it can deduct an additional £75 from its taxable income before calculating its tax liability.

[6] For example, if a company is allowed an enhancement rate of 75% and payable credit rate of 14%, spends £10,000 in R&D, and has no taxable income before the additional deduction, it can claim payable tax credit of $0.14 \times £10{,}000 \times (1 + 0.75) = £2{,}450$. If instead the company has £1,500 in taxable income before the additional deduction, it can first use £2,000 of its R&D to reduce its taxable income to zero (i.e. £1,500 = 75% x £2000, then claim payable tax credit of $0.14 \times £8{,}000 \times (1 + 0.75) = £1{,}960$. This latter case is called a combination claim.

[7] A large company that has taxable loss before the additional deduction therefore may still benefit from R&D tax relief by carrying the "enhanced" loss forward to further reduce its taxable income in the next period. However, this reduction is only meaningful when the company has enough taxable income in this next period.

[8] Qualifying R&D expenditure could include R&D performed outside of the UK by *foreign branches* of UK holding companies, as foreign branches' revenues and costs are directly consolidated into their UK holding companies' tax revenues and costs for UK tax purpose. Qualifying R&D expenditure is unlikely to include R&D performed outside of the UK by *foreign subsidiaries* of UK holding companies, as foreign subsidiaries' net profits are indirectly incorporated into their UK holding companies' tax revenues as dividends for UK tax purpose instead.

expenditure in an accounting period. If an SME works as a subcontractor for a large company, only the subcontractor SME can claim R&D tax relief, under the Large Company Scheme.[9] There is also an upper limit of €7.5m on the total amount of aid a company can receive for any one R&D project under the SME Scheme.[10]

The evolution of the UK R&D Tax Relief Scheme is summarized in Table A1. It was first introduced in April 2000 only for SMEs (Finance Act 2000), then later extended to large companies starting from April 2002 (Finance Act 2002).[11] Between April 2000 and December 2004, the UK followed Recommendation 1996/280/EC for SME definition, which set the ceilings for staff headcount, assets, and sales at 249, €27m, and €40m respectively. From January 2005, the UK adopted Recommendation 2003/361/EC and accordingly increased its SME ceilings to 249, €43m, and €50m. Throughout the period from April 2000 (April 2002) to March 2008, the enhancement rates were set at 50% for SMEs and 25% for large companies, and the payable credit rate for SMEs was 16%.[12]

As discussed in the main paper, the 2007 Finance Act introduced numerous changes to the scheme, in both SME eligibility thresholds and relief rates, which became effective at different points in 2008.[13] First, from April 2008, the enhancement rate for large companies was increased from 25% to 30%. Then from August 2008, the enhancement rate for SMEs was increased from 50% to 75% and the payable credit rate for SMEs was reduced from 16% to 14% (to ensure that state aid intensity stays below 25%). Also from August 2008, the SME Scheme was extended to "larger" SMEs as the SME ceilings were doubled to 499, €86m, and €100m for staff headcount, assets, and sales respectively. This change in SME definition is applicable only for the purpose of the R&D tax relief and therefore is the main focus of our paper, as it allows us to separate the impacts of the R&D Tax Relief Scheme from other programs. It should also be noted that even though these new SME ceilings were announced in Finance Act 2007, the date on which they became effective (August 1st 2008) was appointed much later, in July 2008.[14]

There were tweaks to the system in 2011 and 2012. From April 2011, the SME enhancement rate was increased to 100% and the SME payable credit rate was reduced to 12.5%. From April 2012, the SME enhancement rate was again increased to 125%. However, the SME definition as announced in

[9] An SME already receiving another form of notified state aid for a project cannot claim R&D tax relief for that same project under the SME Scheme (which is also a notified state aid), as total state aid intensity cannot exceed 25% under European Commission's State Aid rules. However, from April 2003 onward, SMEs were allowed to claim R&D tax relief for such projects under the Large Company Scheme.

[10] In practice, most companies prefer R&D Tax Relief's SME Scheme over other notified state aids, as the scheme is sufficiently generous (the maximum relief intensity under the SME Scheme is close to 25%) and is straightforward to apply for. For further details on other conditions, see http://www.hmrc.gov.uk/manuals/cirdmanual/CIRD81000.htm.

[11] Finance Act 2000 (Chapter 17, Schedule 20) and Finance Act 2002 (Chapter 23, Schedule 12).

[12] One exception to this differential treatment of SMEs and large companies was the Vaccine Research Relief Scheme (VRR) launched in April 2003, which extended the higher 50% additional allowance to cover specific areas of vaccine and drug research conducted in large companies (Finance Act 2003, Chapter 14, Schedule 31). The VRR enhancement rate was later reduced to 40% from August 2008 onward.

[13] Finance Act 2007 (Chapter 11).

[14] Finance Act 2007, Section 50 (Appointed Day) Order 2008 of July 16th 2008.

Finance Act 2007 and the large company enhancement rate of 30% remained unchanged throughout this period.

A.3 Model of R&D demand

Consider a CES production function in R&D capital (G) and non-R&D capital (Z). If input markets are competitive we can write the long-run static first order condition for relative factor demand of the firm as:

$$\ln G = -\sigma \ln \rho + \sigma \ln U + \ln Z + B$$

where ρ is the user cost of R&D capital, U is is the user cost of non-R&D capital and B are technological constants representing factor bias terms in the production function. Assume that G can be described by the perpetual inventory formula $G_t = (1-\delta)G_{t-1} + rd_t$ with rd_t the R&D expenditure in period *t*. Since in steady state, the R&D just offsets the depreciated part of the R&D stock $\delta G = rd$, we can re-write the first order condition as:

$$\ln rd = -\sigma \ln \rho + \sigma \ln U + \ln Z + \ln \delta + \mathrm{B}$$

This is essentially the equation we estimate in equation (1). Around the R&D SME threshold the user cost of non-R&D capital and technology are smooth. Non-R&D capital (assets) is the running variable so we have a polynomial approximation to *lnZ*. The only departure from this first-order condition is that we cannot estimate R&D in logarithms because of the presence of firms who don't do any R&D, so we estimate the left hand side in levels instead of logs. To obtain the proportionate change in R&D we use the empirical averages of lagged R&D spending in the pre-policy change period (or explicitly condition on the firm's lagged R&D). We also show that the calculations are robust to using a Poisson regression whose first moment is the exponential log-link function and so is equivalent to estimating in logarithms.

A.4 Tax-adjusted user cost of R&D

The full formula for tax-adjusted user cost of R&D as described in sub-section 5.2 is:

$$\rho_{t,f} = (\Pr(Has\ tax\ liability) \times \frac{\left(1-\tau_t(1+e_{t,f})\right)}{(1-\tau_t)} + \Pr(No\ tax\ liability) \times (1-c_{t,f}(1+e_{t,f}))) \times (r+\delta)$$

where τ is the effective corporate tax rate, e is the enhancement rate, c is the payable credit rate, r is the real interest rate, δ is the depreciation rate, t denotes year, and f denotes the whether the company is an SME or a large company. Note that $\rho_{t,f}$ varies over time with τ_t, $e_{t,f}$, and $c_{t,f}$.

For simplicity, we do not consider the possibility that a loss-making large company may still benefit from R&D tax relief by carrying the "enhanced" loss forward to reduce its taxable income in the next period, as this reduction is only meaningful if the company makes enough profits in this next period. This simplification may overestimate large companies' tax-adjusted user cost of R&D and, as a result,

underestimate R&D tax-price elasticity (by overestimating the difference in tax-adjusted user cost of R&D between SMEs and large companies). We also do not consider combination claims (cases in which an SME combines tax deduction with payable tax credit as described in footnote 6 of sub-section A.2), as there are almost none of them in our baseline sample.

The evolution of tax adjusted user costs of R&D for SMEs and large companies over time is summarized in Table A11. For large companies (for which payable credit rate is always zero), slight decreases in corporate tax rate over 2006-12 (from 30% to 28% to 26%) coupled with slight increases in enhancement rate (from 25% to 30%) over the same period result in a relatively stable tax-adjusted user cost of 0.190 throughout this period. It is therefore reasonable to use the baseline sample's average R&D over 2006-08 as a proxy for how much an average firm in the baseline sample would spend on R&D if it remained a large company over 2009-11, after the policy change. For SMEs, large increases in enhancement rate (from 50% to 75% to 100%) more than offset the slight decrease in corporate tax rate and payable credit rate (from 16% to 14% to 12.5%), leading to a steady reduction in SMEs' tax-adjusted user cost of R&D from 0.154 in 2006 to 0.141 in 2011. This widens the difference in tax-adjusted user cost of R&D between SMEs and large companies over time, from an average log difference of -0.219 over 2006-08 to an average log difference of -0.271 over 2009-11.

Finally, as a robustness check, we also consider using the small profit rate (from 19% to 21% to 20% over 2006-11) instead of the main rate for corporate tax rate. As the tax deduction is less generous with a lower corporate tax rate, the resulting tax-adjusted user cost in the tax deduction case is higher for both SMEs and large companies and their gap is smaller in magnitude (average log difference over 2006-08 is -0.185 and over 2009-11 is -0.229).

A.5 Cost effectiveness analysis of R&D Tax Relief Scheme

A full welfare analysis of the R&D Tax Relief Scheme requires both an analysis of the benefits in terms of (say) the increased GDP generated by the R&D induced by the policy (including spillovers) and the deadweight cost of taxation. We would also need to take a position on other general equilibrium effects such as the increase in the wages of R&D workers due to increased demand (Goolsbee, 1998). As an interim step towards this we follow the convention in the literature which is to calculating a "value for money" ratio $\frac{\Delta_{RD}}{\Delta_{EC}}$ where Δ_{RD} is the amount of R&D induced by the policy and Δ_{EC} is the total amount of additional taxpayer money needed to pay for the scheme (which we call "Exchequer Cost", EC).

A.5.1 2008 extension of the SME Scheme

With respect to the 2008 extension of the SME Scheme to cover "larger" SMEs, Δ_{RD} measures the increase in R&D induced by more generous tax relief under the SME scheme by a firm benefitting from the scheme thanks to the new thresholds. That is, $\Delta_{RD} = RD_{new} - RD_{old}$ where RD_{new} and RD_{old}

are the firm's R&D's under the new and old policies respectively. Similarly, $\Delta_{EC} = EC_{new} - EC_{old}$ where EC_{new} and EC_{old} are the firm's corresponding Exchequer costs due to the policy change.

Rearranging the R&D tax-price elasticity formula $\eta = \frac{\ln(RD_{new}/RD_{old})}{\ln(\rho_{new}/\rho_{old})}$ gives

$$\ln\left(\frac{RD_{new}}{RD_{old}}\right) = \eta \times \ln(\frac{\rho_{new}}{\rho_{old}})$$

where ρ is the tax-adjusted user cost of R&D. For simplicity, we consider the tax deduction case and the payable tax credit case separately.

SME Tax deduction case

In this case,

$$\rho^{deduction} = \frac{\big(1 - \tau(1+e)\big)}{1-\tau}(r+\delta)$$

$$EC^{deduction} = RD \times e \times \tau$$

where τ is the effective corporate tax rate, e is the enhancement rate, r is the real interest rate, and δ is the depreciation rate. As the above firm moves from being a large company pre-2008 to being an SME post-2008, its enhancement rate increases from 25% to 75%. At the same time, corporate tax rate decreases from 30% to 28%. Combining $e_{old} = 0.25, e_{new} = 0.75, \tau_{old} = 0.30, \tau_{new} = 0.28$ with estimated R&D tax-price elasticity of $\eta = -2.63$ gives $\ln\left(\frac{\rho_{new}}{\rho_{old}}\right) = -0.23$ and $\frac{RD_{new}}{RD_{old}} = 1.84$. That is, R&D increases by 84% in response to a 23% reduction in its log user cost.

On the cost side, we have

$$EC_{old} = RD_{old} \times e_{old} \times \tau_{old} = RD_{old} \times 0.075$$

$$EC_{new} = RD_{new} \times e_{new} \times \tau_{new} = RD_{new} \times 0.21.$$

Putting all the elements together gives

$$\frac{\Delta_{RD}}{\Delta_{EC}} = \frac{RD_{new} - RD_{old}}{EC_{new} - EC_{old}} = \frac{(RD_{old} \times 1.84) - RD_{old}}{(RD_{old} \times 1.84 \times 0.21) - (RD_{old} \times 0.075)} = \frac{0.84}{0.31} = 2.70.$$

so the value for money ratio in the tax deduction case is 2.70. In other words, £1 of taxpayer money generates £2.70 in additional R&D.

Finally, note that Δ_{EC} could be rewritten as

$$\Delta_{EC} = EC_{new} - EC_{old} = RD_{new} \times 0.21 - RD_{old} \times 0.075 = \Delta_{RD} \times 0.21 + RD_{old} \times (0.21 - 0.075)$$

where the first element represents the Exchequer costs associated with new R&D and the second term reflects additional Exchequer costs paid on existing R&D due to more generous tax relief. In this case, the majority of the additional costs are because of the new R&D generated, i.e. $\Delta_{RD} \times 0.21 = RD_{old} \times 0.18$ makes up close to 60% of Δ_{EC} ($\Delta_{EC} = RD_{old} \times 0.31$).

SME Payable tax credit case

In this case,

$$\rho^{credit} = (1 - c(1 + e))(r + \delta)$$

$$EC^{credit} = RD \times c \times (1 + e)$$

where c – the payable credit rate – is always zero for large companies and 14% for SMEs post-2008. Combining $c_{old} = 0, c_{new} = 0.14, e_{old} = 0.25, e_{new} = 0.75,$ and $\eta = -2.63$ gives $\ln\frac{\rho_{new}}{\rho_{old}} = -0.28$ and $\frac{RD_{new}}{RD_{old}} = 2.09$ (i.e. R&D increases by 109% in response to a 28% reduction in its log user cost). On the cost side, $EC_{old} = 0$ and $EC_{new} = RD_{new} \times c_{new} \times (1 + e_{new}) = RD_{new} \times 0.25$. Putting all the elements together gives

$$\frac{\Delta_{RD}}{\Delta_{EC}} = \frac{RD_{new} - RD_{old}}{EC_{new} - EC_{old}} = \frac{RD_{old} \times 2.09 - RD_{old}}{RD_{old} \times 2.09 \times 0.25 - 0} = \frac{1.09}{0.51} = 2.13.$$

The value for money ratio in the payable tax credit case is 2.13. In this case, the amount of additional R&D's Exchequer costs due to newly-generated R&D $\Delta_{RD} \times 0.25 = RD_{old} \times 0.27$ constitutes more than 50% of Δ_{EC} ($\Delta_{EC} = RD_{old} \times 0.51$).

A.5.2 R&D Tax Relief Scheme over 2006-11

To evaluate the overall R&D Tax Relief Scheme over 2006-11, we calculate

$$\frac{\Delta_{RD}}{\Delta_{EC}} = \frac{RD_{tax\ relief} - RD_{no\ tax\ relief}}{EC_{tax\ relief} - EC_{no\ tax\ relief}} = \frac{RD_{taxrelief} - RD_{no\ tax\ relief}}{EC}$$

separately for each of three cases, SME tax deduction case (Table A13, Panel B), SME payable tax credit case (Panel C), and large company tax deduction case (Panel D), in each of the years using the same approach as described in detail above. We generalize our estimated tax-price elasticity of 2.6 to the whole population of SMEs, but use a more conservative tax-price elasticity of 1.0 for the population of large companies as these firms are less likely to be credit constrained and therefore less responsive to tax incentives. In addition, we use the small profits rate (19%-21%) instead of the regular corporate tax rate (26%-30%) for the population of SMEs as most of them are much smaller than the "larger" SMEs in our baseline sample and therefore most likely qualify for the small profits rate.

As reported in Table A13, the SME tax deduction's value for money ratio decreases from 2.9 in 2006 to 2.7 in 2011 as SME tax deduction becomes significantly more generous over time. On the other hand, SME payable tax credits and large company tax deduction's value for money ratios are stable at around 2.1 and 1.4 respectively as these schemes do not change much over this period. The fact that all the value for money ratios are well above unity indicates that the R&D Tax Relief Scheme is effective in inducing additional R&D at relatively low cost to the Exchequer.

Finally, we estimate the amount of additional R&D induced by the R&D Tax Relief Scheme as $\Delta_{RD} = \frac{\Delta_{RD}}{\Delta_{EC}} \times EC$ using the calculated value for money ratios $\frac{\Delta_{RD}}{\Delta_{EC}}$ and Exchequer costs national statistics (HMRC 2015). We do this for each of the three schemes in Panels B, C and D and then aggregate them together in Panel D.

To give an example, consider the SME tax deduction case in Panel B for 2009. The user cost is calculated using the policy parameters given, i.e. $\frac{1-0.21\times(1+0.75)}{1-0.21}(0.05+0.15)=0.160$. This compares to a user cost of $0.20 = 0.05 + 0.15$ in the counterfactual world without any tax relief (e = 0). The log difference in user cost is -0.2223. Applying the formula we obtain:

$$\left(\frac{\Delta_{RD}}{\Delta_{EC}}\right)^{deduction} = \frac{1}{e\tau}\left(1 - \frac{1}{\exp(\eta \times \ln(\frac{\rho_{new}}{\rho_{old}}))}\right)$$

Or $(\frac{1}{0.75\times0.21})\ (1 - \frac{1}{\exp(2.63\times0.2223)}) = 2.811$ as in the second row of Table A13 Panel B. From HMRC data we know that £130m was paid out in the SME deduction in this year. Hence, we can calculate that the total amount of additional R&D induced was $£365m = 130 \times 2.811$ (fourth row of Panel B).

As discussed in sub-section 5.4, our estimates suggest that the overall impact of the R&D Tax Relief Scheme is large in Panel E. Over 2006-11, the policy, which costs less than £6 billion in lost tax revenue, induced close to £10 billion in additional R&D. *On an annualized basis, spending £0.96b produced £1.64b of additional R&D.*

These calculations show our estimates of what the counterfactual path of R&D would have been in the absence of the R&D Tax Relief Scheme. The bottom row of Table A13 gives the yearly breakdown. For example, the final column shows that on average 2006-11 we estimate that R&D would be a full 16% lower in the absence of the tax scheme.

It is important to note that throughout our analysis we have been focusing on *qualifying* R&D, i.e. that part of business R&D that is eligible for tax relief. Aggregate qualifying R&D is lower than the figures for Business Enterprise R&D (BERD) reported in Figure 4. For example, in 2011 aggregate BERD was £17bn and aggregate qualifying R&D was £12bn. There are various reasons for this difference, including the fact that BERD includes R&D spending on capital investment whereas qualified R&D does not (only current expenses are liable). It is also the case that HMRC defines R&D more narrowly for tax purposes that BERD which is based on Frascati definition.

We present counterfactual BERD to GDP ratios in Figure 4. To calculate the counterfactual (the dotted line "UK without tax relied" in Figure 4) we simply deduct the additional qualified R&D that we estimate were created by the R&D tax relief system (Table A13, Panel D "Total Additional R&D") from the aggregate BERD numbers from OECD MSTI Dataset (https://stats.oecd.org/Index.aspx?DataSetCode=MSTI_PUB). Since BERD is greater than qualifying R&D, the 16% fall in qualifying R&D translates to about a 10% fall in BERD.

Appendix B: Data

B.1 CT600 dataset

The CT600 dataset is constructed by the UK tax authority (HMRC) and is a confidential panel dataset of corporate tax returns or assessments made from the returns for the universe of companies that file a corporate tax return in the UK. We can only access the dataset from within an HMRC facility (similar to a US Census Bureau Research Data Center) and merging with other datasets requires approval from HMRC. It is currently not possible to merge CT600 with other government secured datasets available at different facilities.[15] The CT600 dataset covers all accounting periods whose end dates fall between April 1st 2001 and March 31st 2012 and consists of all information on the UK Company Tax Return form (which is called the CT600 form). Specifically, an extension of CT600, the Research and Development Tax Credits (RDTC) dataset, provides detailed information on tax relief claims. However, CT600 contains little information on financial statement variables (e.g. assets and employment are not included) as they are not directly required on corporate tax forms.

We convert the original observation unit of firm by accounting period in CT600 to firm by financial year by aggregating all accounting periods the end dates of which fall in the same financial year.[16] This conversion affects a very small number of observations as only 3% of our firm by year observations are aggregates of multiple accounting periods. Our converted dataset then contains 15.7 million firm by year observations over 12 financial years from 2000 to 2011 (covering 3.2 million firms), including 9.1 million firm by year observations over our study period from 2006 to 2011 (covering 2.5 million firms).

Our key variables of interest are those related to firms' R&D tax relief claims from CT600's RDTC dataset, which include the amount of qualifying R&D expenditure each firm has in each year and the scheme under which it makes the claim (SME vs. Large Company Scheme). These variables, originally self-reported by firms on their CT600 forms, have been further validated and corrected by HMRC staff using additional tax processing data available only within the tax authority. It should also be noted that R&D tax relief variables are only available for R&D-tax-relief-claiming firms for the years in which they make the claims. While it is possible to infer that non-claiming firms have zero qualifying R&D expenditure, it is not possible to construct their precise SME eligibility without full information on employment, assets (balance sheet total), sales, and ownership structure.

Table B1 shows that over our study period between 2006 and 2011, we observe claims in 53,491 firm by year observations (by 20,730 firms), 81% of which are under the SME Scheme. The total qualifying R&D expenditure and estimated Exchequer costs under the SME Scheme are in nominal terms

[15] For example, it is currently not possible to merge CT600 with the BERD firm survey which is used to build the national estimate of R&D.

[16] Financial year *t* begins on April 1st of year *t* and ends on March 31st of year *t+1*. So the last year that is currently available to use is 2011.

£11.2bn and £1.8bn respectively; the corresponding figures under the Large Company Scheme are £48.5bn and £3.9bn (excluding claims by SME subcontractors). These figures are in line with the official R&D Tax Relief Scheme statistics released in HMRC (2014).

We also use the data on sales and on investment in plant and machinery from CT600. Sales are reported for 93% of firm by year observations and annualized to account for different accounting period lengths. CT600 tax-accounting sales, which is calculated using the cash-based method, is not the same as financial-accounting sales (reported in the FAME data – see below), which is calculated using the accrual method and used to determine SME eligibility.[17] However, CT600 sales provides a good measure for firms' growth and performance, given its relatively wide coverage.

B.2 FAME dataset

FAME is a database of UK companies provided by Bureau Van Dijk. The panel dataset contains companies' balance sheet and income statement data from companies' annual accounts filed at the UK company registry (Companies House), together with additional information on addresses and industry codes. Like other countries, UK regulations for reporting accounting variables vary with company size, so some balance sheet and income statement variables are missing – we discuss the implications of this below.[18]

Our FAME dataset also covers 12 financial years from 2000 to 2011 and contains 19.6 million firm by year observations (covering 3.8 million firms), including 11.5 million firm by year observations over our study period from 2006-11 (covering 3.1 million firms). Our key SME-eligibility variable from FAME (for R&D tax relief purpose) is total assets (i.e. balance sheet total). As almost all UK companies are required by the Companies House to send in their balance sheets (either full or simplified) for their annual accounts regardless of their size, total assets coverage in FAME is close to complete, at 97% over our study period of 2006-11. On the other hand, sales (financial-accounting sales used to determine SME eligibility) is reported by only 15%, as smaller firms are not required to provide their income statements.[19] The proportion of firms who report employment is even lower at 5%, as employment reporting is not mandatory. Even in our baseline sample of relatively larger firms (i.e. firms with total assets in 2007 between €61m and €111m), the proportion of firms who report sales is 67% and the proportion who report employment is 55%. For this reason, while we do use FAME sales and employment as running variables

[17] The cash-based method focuses on actual cash receipts rather than their related sales transactions. The accrual methods records sale revenues when they are earned, regardless of whether cash from sales has been collected.
[18] All UK limited companies, public limited companies (PLC), and limited liability partnerships (LLP) are required to file *annual accounts* with the Companies House. An annual accounts should generally include a balance sheet, an income statement, a director's report, and an audit report. However, smaller companies may be exempt from sending in income statement, director's report, or audit report. All UK registered companies are required to file *annual returns* with the Companies House, which contain information on registered address and industry codes.
[19] Small companies (those having any 2 of the following: (1) sales of £6.5m or less, (2) assets of £3.26m or less, (3) 50 employees or less) are only required to send in balance sheets. Micro-entities (those having any 2 of the following: (1) sales of £632,000 or less, (2) assets of £316,000 or less, (3) 10 employees or less) are only required to send in simplified balance sheets.

in some alternative specifications, our baseline sample and key results are derived using total assets as the running variable.

Besides total assets, sales, and employment, other FAME variables used in our paper include primary industry code (UK 4-digit SIC), register address postcode, and fixed assets as a proxy for capital stock.

B.3 PATSTAT dataset

Our patent data are drawn from the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office (EPO).[20] PATSTAT is the largest international patent database available to the research community and includes nearly 70 million patent documents from over 60 patent offices, including all of the major offices such as the United States Patent and Trademark office (USPTO), the Japan patent office (JPO) and the Chinese Patent and Trademark Office (SIPO) in addition to the EPO. Therefore, PATSTAT data cover close to the population of all worldwide patents since the 1980s.

PATSTAT reports the name and address of patent applicants, which allows matching individual patents with company databases. The matching between PATSTAT and FAME is implemented by Bureau Van Dijk and is available as part of the ORBIS online platform through a commercial agreement. The quality of the matching is excellent: over our sample period, 94% of patents filed in the UK and 96% of patents filed at the EPO have been matched with their owning company.

A patent is a legal title protecting an invention. To be patented, a product or process must be new, involve an inventive step and be susceptible of industrial application. Patents grant their owner a set of rights of exclusivity over an invention. The legal protection conferred by a patent gives its owner the *right* to exclude others from making, using, selling, offering for sale or importing the patented *invention* for the *term of the patent*, which is usually 20 years from the filing date, and in the country or countries where the patent has been filed (and subsequently granted). In addition to the financial and administrative costs of filing, there is a mandatory public disclosure of the description of the technology, which makes imitation easier and facilitates future technological developments.

To make things clearer, consider a simplified invention process. In the first stage, an inventor discovers a new technology. She then decides where to market this invention and how to protect the intellectual property associated with it. A patent in country *i* grants her an exclusive right to commercially exploit the invention in that country. Accordingly, she will patent her invention in country *i* if she plans to market it there. The set of patents in different countries related to the same invention is called a *patent family*. The vast majority of patent families include only one patent (usually in the home country of the inventor). Importantly, PATSTAT reports not only the unique identifier of each patent application, it also indicates a unique patent family indicator for each patent (we use the DOCDB patent family indicator).

[20] For further details see http://www.epo.org/searching/subscription/raw/product-14-24.html.

This allows us to identify all patent applications filed worldwide by UK-based companies and to avoid double-counting inventions that are protected in several countries.

In this study, our primary measure of innovation is the *number of patent families* – irrespective of where the patents are filed. This proxies for the number of inventions a firm makes. This means that we count the number of patents filed anywhere in the world by firms in our sample, be it at the UK Intellectual Property Office, at the European Patent Office, at the USPTO or anywhere else, but we use information on patent families to make sure that any invention patented in several places is only counted once. Patents are sorted by the first year they were filed (the priority year).

We use fractional counts to account for multiple applicants. For example, if two firms jointly apply for a patent, then each firm is attributed one half of a patent. In practice, 8% of patents filed by UK-based companies are filed jointly by at least two companies.

There are many well-known issues with patents as a measure of innovation. As noted above, not all inventions are patented, although it is reasonable to assume the most valuables are, so counting patents screens out many of the low value inventions. Nevertheless, since patents are of very heterogeneous importance we use several approaches to examine how our results change when looking at patent quality. First, we distinguish between patents filed at the UK patents office and patents files at the EPO. [21] Since the financial and administrative cost is about six times higher at the EPO than UK patent office, EPO patents will, on average be of higher perceived private value. It is also worth noting that the EPO has not experienced the same explosion of low value patents that the US has suffered since the late 1980s (Jaffe and Lerner, 2004)

A second measure of patent quality is the size of patent families, the number of jurisdictions in which each patent is filed. There is evidence that the number of jurisdictions in which a patent is filed is an indicator of its economic value as patenting is costly (see Guellec and van Pottelsberghe, 2000, and Harhoff et al., 2003). A third measure of quality is to distinguish by technology class, as some classes (e.g. pharmaceuticals) are likely to be more valuable than others (e.g. business process methods). Fourth, we use patent citations, also available from PATSTAT. For each patent in the database, we know how many times it was cited by subsequent patents (excluding self-citations). We use the number of subsequent citations (referred to as forward citations) as a measure of value. Again, this measure is well rooted in the patent literature (Squicciarini et al., 2013, Hall et al., 2005, Lanjouw and Schankerman, 2004)

In PATSTAT, patents are categorized based on the International Patent Classification (IPC). We use IPC codes at three-digit level to construct measures of the technological distance between firms used to investigate spillover effects.

[21] Note that because of differences in the "technological scope" of patents across patent offices, two patents filed in the UK may be "merged" into a single patent filed at the EPO. In this case, these three patents will constitute a single patent family and the number of patent families is smaller than the number of UK patents. This configuration happens very rarely, however.

B.4 Sample construction: merging datasets

CT600 was merged with FAME using an HMRC-anonymized version of company registration number (CRN), which is a unique regulatory identifier in both datasets. 95% of CT600 firms between 2006 and 2011 also appear in FAME, covering close to 100% of R&D performing firms and 100% percent of patenting firms in this period.[22] Unmatched firms are slightly smaller but not statistically different from matched ones across different variables reported in CT600, including sales, gross trading profits, and gross and net corporate tax chargeable.[23] Furthermore, that the match rate is less than 100% is due to CRN entering error in FAME, which happens more often among firms that are much smaller than those around SME-eligibility thresholds.[24] For these reasons, we believe sample selection due to incomplete matching between CT600 and FAME is unlikely to be an issue for us.[25]

PATSTAT has been merged with FAME by BVD. As PATSAT comprehensively covers all UK patenting firms, we can safely infer that non-matched firms have zero patents. Over our study period of 2006-11, 9,420 out of 2.5 million CT600 firms claim a total of 46,405 patent families (in 17,293 firm by year observations), including 23,617 higher-quality EPO patents. These patents cover 90% of the total recorded in PATSTAT.

From the merged master dataset, we construct our baseline sample based on total assets in 2007, as it is our key running variable. Specifically, our baseline sample includes 5,888 firms that satisfy the two following conditions: (1) the firm's total assets in 2007 is between €61m and €111m (within €25m below and above the SME threshold of €86m), and (2) the firm appears in CT600 in 2008 or after (to exclude firms exiting before the policy change in 2008). Baseline sample descriptive statistics are summarized in Tables 1 and A2 and discussed in detail in sub-section 4.2.

B.5 Variable construction

As FAME total assets and sales are reported in sterling while the corresponding SME ceilings are set in euros, we convert sterling to euros using the exact same rule used by HMRC for tax purposes. That is, the conversion should be done using the exchange rate on the last day of the relevant accounting period or the average daily exchange rate throughout that accounting period, whichever is more beneficial

[22] Out of 2,495,944 firms present in CT600 between 2006 and 2011, 2,358,948 firms are matched to FAME (94.5% match rate). Over the same period, 20,627 out of 20,730 R&D-performing firms and 9,376 out of 9,420 patenting firms are matched to FAME (99.5% match rate).

[23] Differences (standard errors) between matched and unmatched firms in sales (£'000), gross trading profits (£), gross corporate tax chargeable (£) and net corporate tax chargeable (£) are 970 (3,286), 8,969 (13,703), 3,497 (3,898) and 1,961 (2,291) respectively. None of these differences are statistically significant at conventional level.

[24] Because of confidentiality concerns, we do not get to work directly with CRNs but an anonymized version of CRNs provided by the HMRC Datalab for both FAME and CT600 datasets. This prevents us from further cleaning and matching of initially unmatched firms due to above issue.

[25] The correlation between ln(sales) in CT600 and ln(sales) FAME is 0.90. As noted above, the variables are not measured in the same way, but the fact that their correlation is high is reassuring that the match has been performed well.

for the enterprise. The daily exchange rate is obtained from the OECD, exactly the same method as used by HMRC.

For qualifying R&D expenditure, we do not include the amounts claimed by SME subcontractors, which do not benefit from more generous reliefs under the SME Scheme. Since SME subcontracting makes up only a small portion of the overall R&D Tax Relief Scheme, we believe excluding SME subcontracting does not materially affect our key findings. To account for price differences across years, we also convert nominal values of R&D expenditure to their real values in 2007 price, using UK annual CPI as reported in the World Bank Economic Indicators database.[26]

We address the presence of outliers in R&D spending or patenting by winsorizing our key outcome variables, which include qualifying R&D expenditure and number of all patents as well as number of UK patents, EPO patents and patent families. Specifically, for each variable, the top 2.5% of non-zero values in each year within the sample of firms with 2007 total assets between €46m and €126m are set to the corresponding 97.5 percentile value (i.e. winsorization at 2.5% of non-zero values). This translates into "winsorizing" the R&D of top 5-6 R&D spenders and the number of patents of top 2-4 patenters in the baseline sample in each year. It should be noted that our key findings are robust to alternative choices of winsorization window (e.g. 1% or 5% instead of 2.5%), or to excluding outliers instead of winsorizing outcome variables.

Construction of other variables is generally in the notes to tables. Total Factor Productivity (TFP) is calculated as $\ln(sales) - (1-\propto_l)\ln(capital) - \propto_l \ln(employment)$, where $\propto_l$ is the share of labor costs in total revenue at the two-digit industry level across all firms in the FAME dataset averaged across the 2006-2011 period. Sales are taken from CT600 and capital is fixed assets from FAME. Firm TFP is measured relative to the mean TFP in the two-digit industry.

[26] Ratios of current-£ to 2007-£ derived using UK annual CPI are 1.023 for 2006, 1.000 for 2007, 0.965 for 2008, 0.945 for 2009, 0.915 for 2010, and 0.875 for 2011.

Table A1. Design of UK R&D Tax Relief Scheme, 2000-2012

Effective from		SME ceilings			Enhancement rate		Payable credit rate		Effective for
		Employ-ment	Total assets	Turn-over	SME	Large company	SME	Large company	
2000	April	249	€27m	€40m	50%	0%	16%	0%	Expenditure that incurred on or after April 1st 2000
2002	April	"	"	"	"	25%	"	"	Expenditure that incurred on or after April 1st 2002
2005	January	"	€43m	€50m	"	"	"	"	Accounting period that ended on or after January 1st 2005
2008	April August	499	€86m	€100m	75%	30%	14%	"	Large companies: expenditure that incurred on or after April 1st 2008 SMEs: expenditure that incurred on or after August 1st 2008
2011	April	"	"	"	100%	"	12.5%	"	Expenditure that incurred on or after April 1st 2011
2012	April	"	"	"	125%	"	"	"	Expenditure that incurred on or after April 1st 2012

Note: To be considered an SME, a company must fall below the employment ceiling and either the total asset ceiling or the sales ceiling ("ceiling tests"). The measurements and account aggregation rules for employment, total assets, and sales are set according to 1996/280/EC (up to December 31st 2004) and 2003/361/EC (from January 1st 2005). A company loses (acquires) its SME status if it fails (passes) the ceiling tests over two consecutive accounting periods (two-year rule). An SME working as subcontractor for a large company can only claim under the Large Company Scheme. From April 2000 to March 2012, there was a minimum requirement of £10,000 in qualifying R&D expenditure for both SMEs and large companies.

Table A2. Baseline sample descriptive statistics, before and after policy change

Subsample	Firms with total assets in 2007 b/w €61-86 million		Firms with total assets in 2007 b/w €86-111 million		Difference b/w firms with total assets in 2007 below vs. above €86 million	
Year	2006-2008	2009-2011	2006-2008	2009-2011	2006-2008	2009-2010
No. firms with qual. R&D exp.	140	160	84	94		
Avg. qual. R&D exp. *(£ '000)*	4,413	4,807	7,850	6,954	-3,437	-2,147
No. firms with patents	104	99	66	58		
Avg. patents	6.29	6.26	6.18	5.03	0.11	1.23
No. firms with UK patents	95	91	58	53		
Avg. UK patents	8.97	8.18	8.70	6.66	0.27	1.52
No. firms with EPO patents	72	54	44	37		
Avg. EPO patents	4.77	5.52	4.82	4.35	-0.05	1.17
Total no. firms in the subsample	3,561		2,327			
Avg. qual. R&D exp. *(£ '000)*	173.5	216.0	283.4	280.9	-109.9	-64.9
Avg. patents	0.184	0.174	0.175	0.125	0.009	0.049
Avg. UK patents	0.239	0.209	0.217	0.152	0.022	0.057
Avg. EPO patents	0.097	0.084	0.091	0.069	0.006	0.015

Note: The baseline sample includes 5,888 firms with total assets in 2007 between €61m and €111m. Total assets data come from FAME and are converted to € from £ using HMRC rule. Qualifying R&D expenditure comes from CT600 panel dataset and are converted to 2007 prices. Patent counts come from PATSTAT.

Table A3. Robustness checks for R&D regressions

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**R&D expenditure, 2009-11 average *(£ '000)***									
Specification	Pooling 2009-11	Higher order polynomial controls		Lagged dependent variable controls (LDV)		Industry & location fixed effects			Poisson specification	
Below new SME asset threshold in 2007	138.5**	171.2*	175.3	75.5**	82.0**	125.6**	107.4**	62.3	1.62***	1.08***
	(55.3)	(87.4)	(108.0)	(37.6)	(36.4)	(61.2)	(49.5)	(38.4)	(0.57)	(0.54)
Past qualifying R&D expenditure (*£ '000)*				0.66***	0.89***					
				(0.08)	(0.09)					
Polynomial controls	1st order	2nd order	3rd order	1st order	1st order	1st order	1st order	1st order	1st order	1st order
Year of LDV				2007	2006-08					2007
Fixed effects	Year					Industry	Location	Ind. x loc.		
Firms	17,664	5,888	5,888	5,888	5,888	4,502	5,768	4,466	5,888	5,888

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	**R&D expenditure, 2009-11 average *(£ '000)***								
Specification	Alternative bandwidth around the threshold				Alternative kernel weight		Alternative winsorization parameter		
Below asset dummy threshold (in 2007)	43.6	77.9*	143.2**	182**	148.6**	151.9**	171.1**	103.2**	210.4**
	(43.5)	(46.9)	(58.4)	(73.5)	(57.8)	(60.8)	(68.1)	(42.2)	(88.9)
Sample total assets	51-121m	56-116m	66-106m	71-101m	61-111m	61-111m	61-111m	61-111m	61-111m
Kernel weight					Epa	Tri			
Winsorization window	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%	1.0%	5.0%	no outliers
Firms	8,818	7,255	4,615	3,384	5,888	5,888	5,888	5,888	5,884

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. Robust standard errors are in brackets. **Panel A:** Column 1 pools observations across 2009-11 with year dummies and standard errors clustered at firm level. Columns 2-3 control for second or third order polynomials of running variable. Columns 4-5 add lagged dependent variable controls. Columns 6-8 add industry (4-digit SIC), location (2-digit postcode), and industry x location (2-digit SIC x 1-digit postcode) fixed effects. Columns 9-10 use Poisson specification instead of OLS, without (column 9) and with (column 10) lagged dependent variable control. **Panel B:** Columns 1-4 use samples with different sample bandwidths around the threshold. Columns 5-6 use Epanechnikov or triangular kernel weights. Columns 7-9 use samples with different winsorization parameters or sample excluding outliers in R&D expenditure.

Table A4. Robustness checks for reduced-form patent regressions

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**All patent count, 2009-11 average**									
Specification	Pooling 2009-11	Higher order polynomial controls		Lagged dependent variable controls (LDV)		Industry & location fixed effects			Poisson specification	
Below new SME asset threshold in 2007	0.073***	0.067	0.067	0.041**	0.043**	0.069*	0.075***	0.092***	1.52***	1.42**
	(0.026)	(0.046)	(0.054)	(0.021)	(0.018)	(0.035)	(0.027)	(0.031)	(0.50)	(0.57)
Past qualifying R&D expenditure (*£ '000)*				0.738***	0.811***					
				(0.109)	(0.010)					
Polynomial controls	1st order	2nd order	3rd order	1st order	1st order	1st order	1st order	1st order	1st order	1st order
Year of LDV				2007	2006-08					2007
Fixed effects	Year					Industry	Location	Ind. x loc.		
Firms	17,664	5,888	5,888	5,888	5,888	4,502	5,768	4,466	5,888	5,888

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	**All patent count, 2009-11 average**								
Specification	Alternative bandwidth around the threshold				Alternative kernel weight		Alternative winsorization parameter		
Below asset threshold dummy (in 2007)	0.045*	0.076***	0.067**	0.071	0.071***	0.07**	0.07**	0.073***	0.074***
	(0.025)	(0.027)	(0.029)	(0.049)	(0.027)	(0.029)	(0.027)	(0.025)	(0.025)
Sample total assets	51-121m	56-116m	66-106m	71-101m	61-111m	61-111m	61-111m	61-111m	61-111m
Kernel weight					Epa	Tri			
Winsorization window	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%	1.0%	5.0%	no outliers
Firms	8,818	7,255	4,615	3,384	5,888	5,888	5,888	5,888	5,884

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. Robust standard errors are in brackets. **Panel A:** Column 1 pools observations across 2009-11 with year fixed effects and standard errors clustered at firm level. Columns 2-3 control for second or third order polynomials of running variable. Columns 4-5 add lagged dependent variable controls. Columns 6-8 add industry (4-digit SIC), location (2-digit postcode), and industry x location (2-digit SIC x 1-digit postcode) fixed effects. Columns 9-10 use Poisson specification instead of OLS, without (column 9) and with (column 10) lagged dependent variable control. **Panel B:** Columns 1-4 use samples with different sample bandwidths around the threshold. Columns 5-6 use Epanechnikov or triangular kernel weights. Columns 7-9 use samples with different winsorization parameters or sample excluding outliers in R&D expenditure.

Table A5. Robustness checks for effects of R&D Tax Relief Scheme on patents (IV regressions)

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	**Total all patents filed over 2009-11**							
Specification	Pooling 2009-11	Higher order polynomial controls		Lagged dependent variable controls (LDV)		Industry & location fixed effects		
Qual. R&D expenditure over 2009-11 *(£ mill)*	0.530**	0.391	0.382	0.331	0.335*	0.549	0.702**	1.48*
	(0.254)	(0.292)	(0.336)	(0.208)	(0.194)	(0.360)	(0.354)	(0.884)
Past qualifying R&D expenditure (*£ mill)*				0.635***	0.708***			
				(0.125)	(0.122)			
Polynomial controls	1st order	2nd order	3rd order	1st order	1st order	1st order	1st order	1st order
Year of LDV				2007	2006-08			
Fixed effects	Year					Industry	Location	Ind. x loc.
Observations	17,664	5,888	5,888	5,888	5,888	4,502	5,768	4,466

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	**All patents filed, 2009-11 average**								
Specification	Alternative bandwidth around the threshold				Alternative kernel weight		Alternative winsorization parameter		
R&D exp. *(£ mill)*, 2009-11 average	1.03	0.978	0.465*	0.388	0.478**	0.464**	0.429**	0.711**	0.350*
	(1.07)	(0.625)	(0.250)	(0.270)	(0.234)	(0.235)	(0.210)	(0.335)	(0.179)
Sample total assets	51-121m	56-116m	66-106m	71-101m	61-111m	61-111m	61-111m	61-111m	61-111m
Kernel weight					Epa	Tri			
Winsorization window	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%	1.0%	5.0%	no outliers
Firms	8,818	7,255	4,615	3,384	5,888	5,888	5,888	5,888	5,884

Note: *** significant at 1% level, ** 5% level, * 10% level. IV estimates based on the (fuzzy) RD design. Instrumental variable is the dummy whether total assets in 2007 is below €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of the running variable (total assets in 2007) separately for each side of the threshold are included. Robust standard errors are in brackets. **Panel A:** Column 1 pools observations across 2009-11 with year fixed effects and standard errors clustered at firm level. Columns 2-3 control for second or third order polynomials of running variable. Columns 4-5 add lagged dependent variable controls. Columns 6-8 add industry (4-digit SIC), location (2-digit postcode), and industry x location (2-digit SIC x 1-digit postcode) fixed effects. **Panel B:** Columns 1-4 use samples with different sample bandwidths around the threshold. Columns 5-6 use Epanechnikov or triangular kernel weights. Columns 7-9 use samples with different winsorization parameters or sample excluding outliers in R&D expenditure.

Table A6. Discontinuities in the probabilities of doing any R&D or filing any patents

Year	(1)	(2)	(3)	(4)	(5)	(6)
	Before (pre-policy)			After (post-policy)		
	2006	2007	2008	2009	2010	2011
Dependent variable	**Dummy: R&D expenditure > 0**					
Below asset threshold dummy (in 2007)	0.011	0.016*	-0.0045	0.007	0.002	0.010
	(0.008)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)
Mean over 2006-08	*0.021*	*0.026*	*0.029*	*0.028*	*0.030*	*0.031*
Dependent variable	**Dummy: All patent count > 0**					
Below asset threshold dummy (in 2007)	0.006	0.009	0.0040	0.011*	0.008	0.012*
	(0.008)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
Mean over 2006-08	*0.021*	*0.020*	*0.015*	*0.016*	*0.017*	*0.015*
Firms	5,888	5,888	5,888	5,888	5,888	5,888

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Dependent variables are dummies indicating whether a firm has R&D expenditure or files patents during the corresponding year.

Table A7. Heterogeneous effects of R&D Tax Relief Scheme by past R&D and patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Specification	First stage OLS		Reduced form OLS					
Dependent variable (2009-11 average)	**R&D expenditure *(£ '000)***		**All patents counts**		**UK patent counts**		**EPO patent counts**	
Subsample	Past R&D > 0	Past R&D = 0	Past all pat. > 0	Past all pat. = 0	Past UK pat. > 0	Past UK pat. = 0	Past EPO pat. > 0	Past EPO pat. = 0
Below asset threshold dummy (in 2007)	2,775**	0.0	1.80***	0.00	2.52***	0.00	1.58**	0.00
	(1,134)	(7.1)	(0.66)	(0.00)	(0.91)	(0.01)	(0.61)	(0.00)
Mean over 2006-08	1,901	0.0	2.08	0.00	2.96	0.00	1.60	0.00
Difference between having vs. not having R&D/patents	2,775		1.80***		2.52***		1.58**	
	(1,125)		(0.65)		(0.90)		(0.60)	
Firms	224	5,664	170	5,718	153	5,735	116	5,772

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. Robust standard errors are in brackets.

Table A8. Heterogeneous effects of R&D Tax Relief Scheme by industry patenting intensity

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Specification	First stage OLS		Reduced form OLS					
Dependent variable (2009-11 average)	**R&D expenditure *(£ '000)***		**All patent count**		**UK patent count**		**EPO patent count**	
Subsample	High patent	Low patent	High patent	Low patent	High patent	Low patent	High patent	Low patent
Below asset threshold dummy (in 2007)	204.6*	100.6	0.16**	0.02	0.21**	0.02	0.08*	0.01
	(106.3)	(67.9)	(0.06)	(0.01)	(0.08)	(0.02)	(0.04)	(0.01)
Mean over 2006-08	117.0	22.3	0.12	0.01	0.15	0.02	0.06	0.01
Difference between high vs. low patenting industries	104.0		0.14**		0.19**		0.06	
	(126.1)		(0.07)		(0.08)		(0.04)	
Firms	2,273	2,231	2,273	2,231	2,273	2,231	2,273	2,231

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)
Specification	IV					
Dependent variable (2009-11 average)	**All patent count**		**UK patent count**		**EPO patent count**	
Subsample	High patent	Low patent	High patent	Low patent	High patent	Low patent
R&D expenditure *(£ million)*, 2009-11 average	0.803*	0.198	1.03*	0.187	0.374	0.119
	(0.478)	(0.161)	(0.604)	(0.199)	(0.249)	(0.103)
Firms	2,273	2,231	2,273	2,231	2,273	2,231

Note: *** significant at 1% level, ** 5% level, * 10% level. Robust standard errors are in brackets. Industry patenting intensity is calculated as the share of firms in the industry (at the 4-digit SIC level) having filed any patent before 2007. **Panel A:** OLS estimates based on the RD design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. **Panel B:** IV estimates based on the (fuzzy) RD design. Instrumental variable is the dummy whether total assets in 2007 is below €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of RDD running variable (total assets in 2007) separately for each side of the threshold are included.

Table A9. Heterogeneous effects of R&D Tax Relief Scheme by firms' past capital investments

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Specification	First stage OLS		Reduced form OLS					
Dependent variable (2009-11 average)	**R&D expenditure *(£ '000)***		**All patent count**		**UK patent count**		**EPO patent count**	
Sample industries	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0
Below asset threshold dummy (in 2007)	338.3***	-37.7	0.16***	0.00	0.21***	0.00	0.08**	0.00
	(113.9)	(32.2)	(0.05)	(0.02)	(0.07)	(0.02)	(0.03)	(0.01)
Mean over 2006-08	153.9	5.6	0.12	0.01	0.15	0.02	0.06	0.01
Difference between high vs. low investment firms	376.0***		0.16***		0.21***		0.08**	
	(118.4)		(0.06)		(0.07)		(0.04)	
Firms	2,655	3,042	2,655	3,042	2,655	3,042	2,655	3,042

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)
Specification	IV					
Dependent variable (2009-11 average)	**All patent count**		**UK patent count**		**EPO patent count**	
Subsample	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0
R&D expenditure *(£ million)*, 2009-11 average	0.470**	0.034	0.611**	0.085	0.237**	0.037
	(0.202)	(0.396)	(0.262)	(0.443)	(0.113)	(0.215)
Firms	2,655	3,042	2,655	3,042	2,655	3,042

Note: *** significant at 1% level, ** 5% level, * 10% level. Robust standard errors are in brackets. Past capital investments is calculated as average machinery and plant investments over 2005-07. **Panel A:** OLS estimates based on the RD design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. **Panel B:** IV estimates based on the RD design. Instrumental variable is the dummy whether total assets in 2007 is below €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of RDD running variable (total assets in 2007) separately for each side of the threshold are included.

Table A10. Estimating impacts of R&D Tax Relief Scheme using other SME criteria

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SME criterion	Total assets		Sales				Employment	
Dependent variable (2009-2011 average)	**R&D exp. *(£ '000)***	**All patent count**	**R&D exp. *(£ '000)***	**All patent count**	**R&D exp. *(£ '000)***	**All patent count**	**R&D exp. *(£ '000)***	**All patent count**
Below SME threshold dummy (in 2007)	138.5**	0.073***	133.9**	0.035	133.0	0.109	77.2	0.120*
	(55.3)	(0.026)	(66.5)	(0.050)	(129.6)	(0.071)	(114.3)	(0.062)
Mean over 2006-08	*72.3*	*0.060*	*105.2*	*0.083*	*176.9*	*0.114*	*197.6*	*0.141*
Treatment effect to baseline ratio	*1.92*	*1.22*	*1.27*	*0.42*	*0.75*	*0.96*	*0.39*	*0.85*
Sample	Total assets in [€61m, €111m]		Sales in [€50m, €150m]		Sales in [€50m, €150m] & total assets > €86m		Employment in [300, 700]	
Firms	5,888	5,888	7,101	7,101	2,085	2,085	4,526	4,526

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)
Specification	First stage	Reduced form	IV	First stage	Reduced form	IV
Dependent variable (2009-2011 average)	**R&D exp. *(£ '000)***	**All patent count**	**All patent count**	**R&D exp. *(£ '000)***	**All patent count**	**All patent count**
Below asset threshold dummy (in 2007)	87.3	0.114***		73.6*	0.079***	
	(59.1)	(0.042)		(41.3)	(0.026)	
Below sales threshold dummy (in 2007)	126.5*	0.032		86.0**	-0.005	
	(66.4)	(0.050)		(43.0)	(0.024)	
R&D expenditure (*£ million)*, 2009-11 average			0.698*			0.410
			(0.405)			(0.238)
Mean over 2006-08	*105.3*	*0.083*	*0.083*	*98.1*	*0.071*	*0.071*
Joint F-statistics (p-value)	2.43 (0.09)	4.04 (0.02)		2.52 (0.08)	5.47 (0.00)	
Sample	Sales in [€50m, €150m]			Total assets in [€61m, €111m] or sales in [€60m, €140m]		
Firms	7,091	7,091	7,091	8,120	8,120	8,120

Note: *** significant at 1% level, ** 5% level, * 10% level. Robust standard errors are in brackets. **Panel A:** OLS estimates based on the RD design. The running variable in columns 1-2 is total assets in 2007 with threshold of €86m. The running variable in columns 3-6 is sales in 2007 with threshold of €100m. The running variable in columns 7-8 is employment in 2007 with threshold of 500. Controls for first order polynomials of running variable separately for each side of the threshold are included. **Panel B:** OLS estimates based on the RD design for first-stage and reduced-form regressions (columns 1-2 and 4-5). IV estimates based on the (fuzzy) RD design where the instrumental variable is the dummy whether total assets in 2007 is below €86m (columns 3 and 6). The running variables are total assets in 2007 with threshold of €86m and sales in 2007 with threshold of €100m. Instrumental variable in columns 3 and 6 are the dummy whether total assets in 2007 is below €86m and the dummy whether sales in 2007 is below €100m. Controls for first order polynomials of the running variable (total assets in 2007 and sales in 2007) separately for each side of the respective threshold are included. Reported joint F-statistics for are for below-asset-threshold dummy and below-sales-threshold dummy.

Table A11. Tax-adjusted user cost of R&D capital over time

Tax relief scheme	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	SME			Large company			Log diff. in user cost
	Deduction	Payable credit	Average	Deduction	Payable credit	Average	
2006	0.157	0.152	0.154	0.179	0.200	0.190	-0.210
2007	0.157	0.152	0.154	0.179	0.200	0.190	-0.210
2008	0.147	0.151	0.149	0.177	0.200	0.190	-0.238
2009	0.142	0.151	0.147	0.177	0.200	0.190	-0.255
2010	0.142	0.151	0.147	0.177	0.200	0.190	-0.255
2011	0.130	0.150	0.141	0.179	0.200	0.191	-0.302
2006-2008	0.154	0.152	0.153	0.178	0.200	0.190	-0.219
2009-2011	**0.138**	**0.151**	**0.145**	**0.177**	**0.200**	**0.190**	**-0.271**

Note: Tax-adjusted user cost of R&D capital is calculated using formulae as described in sub-section 5.2. Corporate tax rate is 30% in 2006-07, 28% in 2008-10, and 26% in 2011. Enhancement rate is 50% for SMEs and 25% for large companies in 2006-08, 75% for SMEs and 30% for large companies in 2008-10, 100% for SMEs and 30% for large companies in 2011. Payable credit rate is 16% in 2006-08, 14% in 2008-10, and 12.5% in 2011. Share of the payable credit case is 55%. Real interest rate is 5%. Depreciation rate is 15%.

Table A12. Tax-price elasticity of R&D investments using different approaches

	Approach	(1) Treatment effect	(2) Baseline R&D	(3) Log diff. in R&D	(4) Log diff. in user cost	(5) **Tax-price elasticity**
1	**Treatment effect:** discontinuity in increase in average R&D expenditure 3-year pre- compared to 3-year post-policy change *Difference(After - Before)* **Baseline R&D:** average R&D expenditure 3-year pre-policy change 2006-08	£75k	£72k	0.71	-0.27	**-2.63**
2	**Treatment effect:** discontinuity in increase in average R&D expenditure 3-year pre- compared to 3-year post-policy change *Difference(After - Before)* **Baseline R&D:** average R&D expenditure 2-year pre-policy change 2006-07	£75k	£75k	0.69	-0.27	**-2.56**
3	**Treatment effect:** discontinuity in average R&D expenditure over 2009-11, controlling for R&D in 2007 **Baseline R&D:** predicted average R&D expenditure over 2009-11 by a large company at the asset threshold €86m	£75k	£47k	0.95	-0.27	**-3.51**
4	**Specification:** Poisson regression **Treatment effect:** discontinuity in average R&D expenditure over 2009-11, controlling for R&D in 2007			1.08	-0.27	**-3.97**
5	**Sample:** firms with total assets in 2007 in [€66m, €106m] **Treatment effect:** discontinuity in increase in average R&D expenditure 3-year pre- compared to 3-year post-policy change *Difference(After - Before)* **Baseline R&D:** average R&D expenditure 3-year pre-policy change 2006-08	£288k	£73k	0.79	-0.27	**-2.91**
6	**Treatment effect:** discontinuity in increase in average R&D expenditure 3-year pre- compared to 3-year post-policy change *Difference(After - Before)* **Baseline R&D:** average R&D expenditure 3-year pre-policy change 2006-08 **Tax-adjusted user cost of R&D capital:** calculated using small profit rate instead of main rate for corporate tax rate	£75k	£72k	0.71	-0.23	**-3.11**

Note: Log difference in R&D investments is calculated as $\ln(treatment\ effect + baseline\ R\&D) - \ln(baseline\ R\&D)$. Tax-price elasticity of R&D is calculated as $\frac{\ln(difference\ in\ R\&D\ investments)}{\ln(difference\ in\ user\ cost\ of\ R\&D\ capital)}$. Baseline sample includes firms with total assets in 2007 in in [€61m, €111m] unless indicated otherwise. Treatment effect used in approaches 1, 2 and 6 is reported in column 9 of Table 2. Treatment effects used in approaches 3 and 4 are reported in columns 4 and 9 of Table A3 Panel A respectively. Treatment effect used in approach 5 is estimated using same specification as in column 9 of Table 2 for the specified sample. Approaches 1 to 5 use the baseline log difference in tax-adjusted user cost of R&D capital between SMEs and large companies as estimated in sub-section 5.2 and reported in Table A12. Approach 6 uses the same formulae as described in sub-section 5.2 to calculate log difference in tax-adjusted user cost of R&D capital, but using small profit rate (20% in 2006-07, 21% in 2008-10, and 20% in 2011) instead of main rate for corporate tax rate.

Table A13. Value for money analysis of R&D Tax Relief Scheme

	2006	2007	2008	2009	2010	2011	Average 2006-11
Panel A: Policy parameters							
SME enhancement rate e_{SME}	50%	50%	67%	75%	75%	100%	
SME payable credit rate c_{SME}	16%	16%	15%	14%	14%	12.5%	
SME effective corporate tax rate τ_{SME}	19%	19%	21%	21%	21%	20%	
LCO enhancement rate e_{LCO}	25%	25%	30%	30%	30%	30%	
LCO effective corporate tax rate τ_{LCO}	30%	30%	28%	28%	28%	26%	
Panel B: SME tax deduction case							
Tax-adjusted user cost of R&D	0.177	0.177	0.165	0.160	0.160	0.150	
Value for money ratio Δ_{RD}/Δ_{EC}	2.944	2.944	2.866	2.811	2.811	2.654	2.791
Exchequer costs *(£m)*	50	60	80	130	160	210	115
Additional R&D *(£m)*	147	177	229	365	450	557	321
Panel C: SME payable tax credit case							
Tax-adjusted user cost of R&D	0.152	0.152	0.151	0.151	0.151	0.150	
Value for money ratio Δ_{RD}/Δ_{EC}	2.142	2.142	2.134	2.133	2.133	2.123	2.134
Exchequer costs *(£m)*	150	180	190	190	190	220	187
Additional R&D *(£m)*	321	386	405	405	405	467	398
Panel D: Large company tax deduction case							
Tax-adjusted user cost of R&D	0.179	0.179	0.177	0.177	0.177	0.179	
Value for money ratio Δ_{RD}/Δ_{EC}	1.429	1.429	1.389	1.389	1.389	1.351	1.392
Exchequer costs *(£m)*	480	550	730	670	750	780	660
Additional R&D by LCOs *(£m)*	686	786	1,014	931	1,042	1,054	919
Panel D: Aggregates							
Total Exchequer costs *(£m)*	680	790	1,000	990	1,100	1,210	962
Total additional R&D *(£m)*	1,154	1,348	1,649	1,701	1,897	2,078	1,638
Value for money ratio	1.697	1.706	1.649	1.718	1.724	1.718	1.703
Total qualifying R&D *(£m)*	7,670	8,880	10,800	9,730	10,880	11,840	9,967
Fall of aggregate R&D without policy	15%	15%	15%	17%	17%	18%	16%

Note: Tax-adjusted user cost of R&D and value for money ratio are calculated using the formulae as described in Appendix A5 using the above policy parameters. In addition, real interest rate is 5% and depreciation rate is 15%. Tax-adjusted user cost of R&D without any tax relief is calculated to be 0.200. Tax-price elasticity of R&D among SMEs is -2.63 as estimated in sub-section 5.2. Tax-price elasticity of R&D among large companies is -1.00. Exchequer costs come from HMRC national statistics. Additional R&D is calculated as value for money ratios times Exchequer costs.

Table B1 Descriptive Statistics

Panel A. Full CT600 dataset

	Unit	**2006**	**2007**	**2008**	**2009**	**20210**	**2011**	**2006-11**
No. of firms	*Firm*	1,406,696	1,487,173	1,484,311	1,504,927	1,564,871	1,646,641	2,495,944
No. of firms claiming R&D relief	*Firm*	6,431	7,429	8,334	9,144	10,150	12,003	20,730
SME Scheme								
No. of firms claiming	*Firm*	5,153	5,855	6,570	7,354	8,238	9,921	-
Avg. qual. R&D expenditure	*£ (nom)*	257,752	268,904	266,730	244,854	263,811	258,541	-
Avg. estimated Exchequer costs	*£ (nom)*	39,433	42,150	41,018	44,099	43,138	43,451	-
Large Company Scheme								
No. of firms claiming	*Firm*	1,290	1,592	1,776	1,795	1,923	2,092	-
Avg. qual. R&D expenditure	*£ (nom)*	4,926,939	4,616,811	5,120,979	4,435,308	4,508,202	4,357,442	-
Avg. estimated Exchequer costs	*£ (nom)*	371,097	346,616	412,088	376,405	382,284	357,870	-
SME subcontractors								
No. of firms claiming	*Firm*	399	443	522	610	720	715	-
Avg. qual. R&D expenditure	*£ (nom)*	630,098	465,590	406,302	504,624	658,942	928,208	-
Avg. estimated Exchequer costs	*£ (nom)*	47,406	48,014	43,043	42,618	46,771	56,809	-
Patenting								
No. of firms having patents	*Firm*	3,093	3,085	2,965	2,806	2,682	2,662	9,420
Avg. number of patents	*Patent*	2.68	2.77	2.72	2.63	2.66	2.64	4.93
No. of firms having UK patents	*Firm*	3,262	3,316	3,228	3,083	2,989	2,965	-
Avg. number of UK patents	*Patent*	3.00	3.08	3.00	2.83	2.78	2.82	-
No. of firms having EPO patents	*Firm*	1,453	1,448	1,376	1,409	1,358	1,125	-
Avg. number of EPO patents	*Patent*	0.95	0.90	0.82	0.83	0.47	0.17	-

Panel B. Full FAME dataset

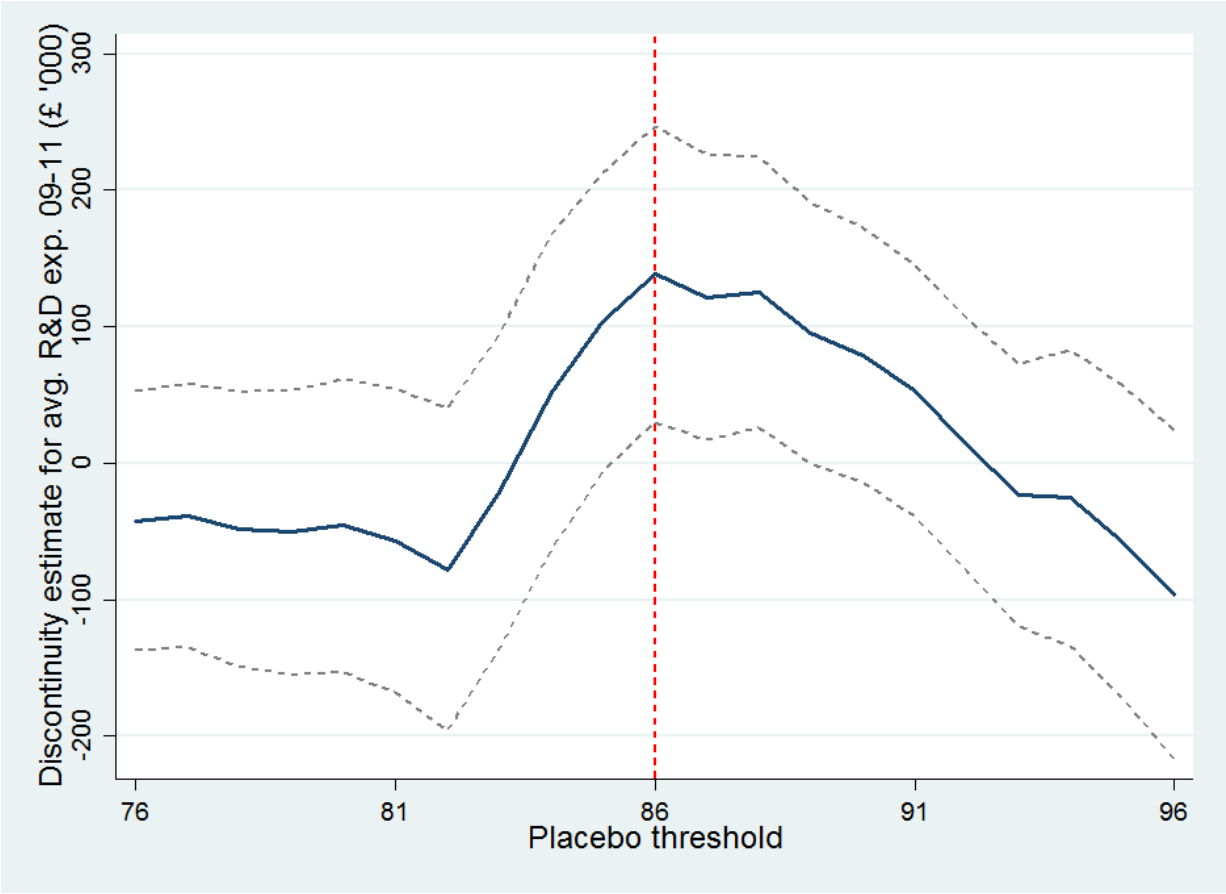
	Unit	**2006**	**2007**	**2008**	**2009**	**20210**	**2011**	**2006-11**
No. of firms	*Firm*	1,780,531	1,858,209	1,870,089	1,898,721	1,973,722	2,073,930	3,140,060
Variable coverage								
No. of firms with total assets	*Firm*	1,732,169	1,807,743	1,818,448	1,843,896	1,914,848	2,015,058	-
Total assets coverage	*%*	97.3%	97.3%	97.2%	97.1%	97.0%	97.2%	-
No. of firms with sales	*Firm*	352,680	319,726	275,938	274,768	263,394	227,463	-
Sales coverage	*%*	19.8%	17.2%	14.8%	14.5%	13.3%	11.0%	-
No. of firms with employment	*Firm*	95,615	93,855	91,375	94,332	98,426	97,814	-
Employment coverage	*%*	5.4%	5.1%	4.9%	5.0%	5.0%	4.7%	-

Panel C. CT600 and FAME matching

	Unit	**2006**	**2007**	**2008**	**2009**	**20210**	**2011**	**2006-11**
# of CT600 firms that appear in FAME between 2006 and 2011	*Firm*	1,353,844	1,427,132	1,442,619	1,468,000	1,529,317	1,598,012	2,358,948
As share of CT600 firm count	*%*	96.2%	96.0%	97.2%	97.5%	97.7%	97.0%	94.5%
Out of which								
# of firms claiming tax relief	*Firm*	6,411	7,409	8,298	9,105	10,108	11,937	20,627
As share of CT600 R&D firms	*%*	99.7%	99.7%	99.6%	99.6%	99.6%	99.5%	99.5%
# of firms having patents	*Firm*	3,078	3,065	2,951	2,789	2,665	2,634	9,376
As share of CT600 patenting firms	*%*	99.5%	99.4%	99.5%	99.4%	99.4%	98.9%	99.5%

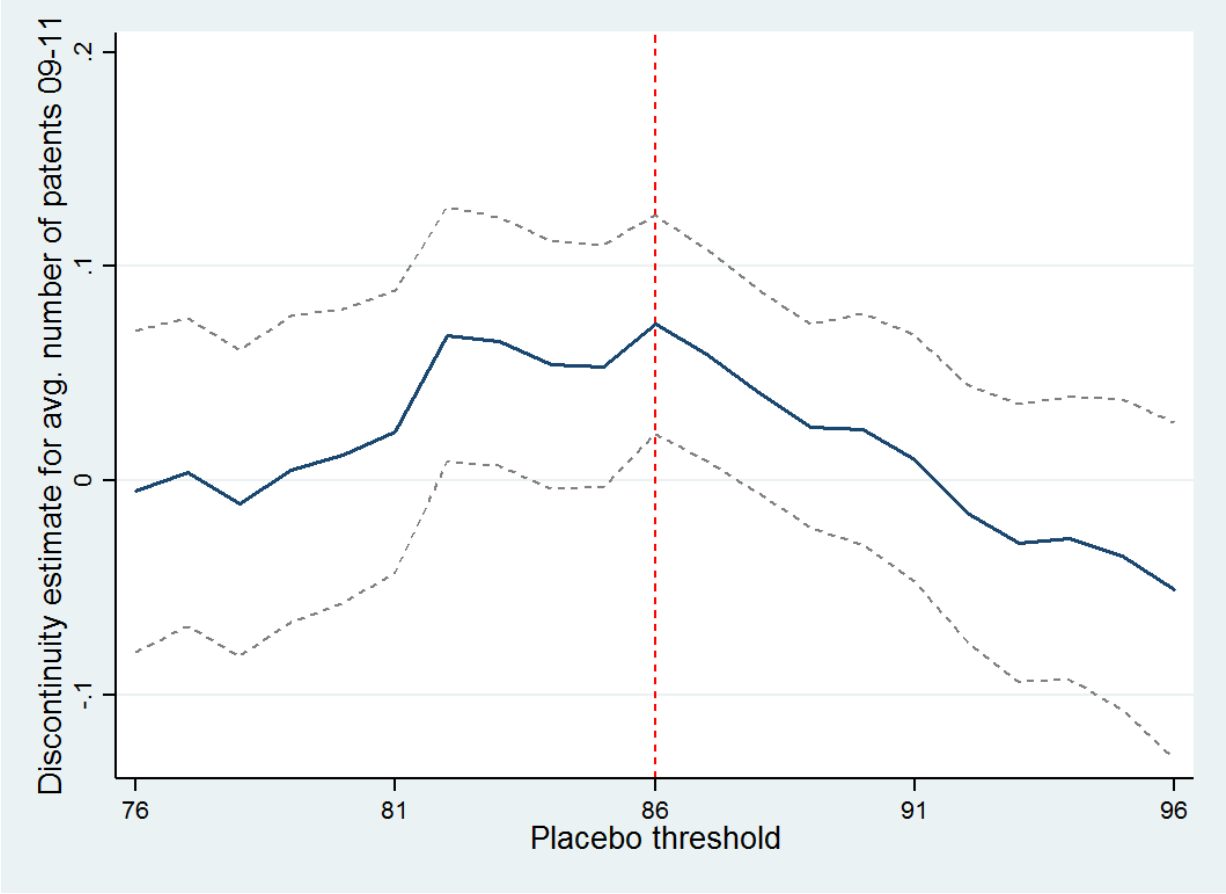
Note: Average qualifying R&D expenditure and estimated Exchequer costs are calculated for corresponding R&D-tax-relief claiming firms. Average number of patents, UK patents, and EPO patents are calculated for corresponding patenting firms.

Figure A1. Discontinuities in average R&D expenditure over 2009-11 at "pseudo" SME asset thresholds



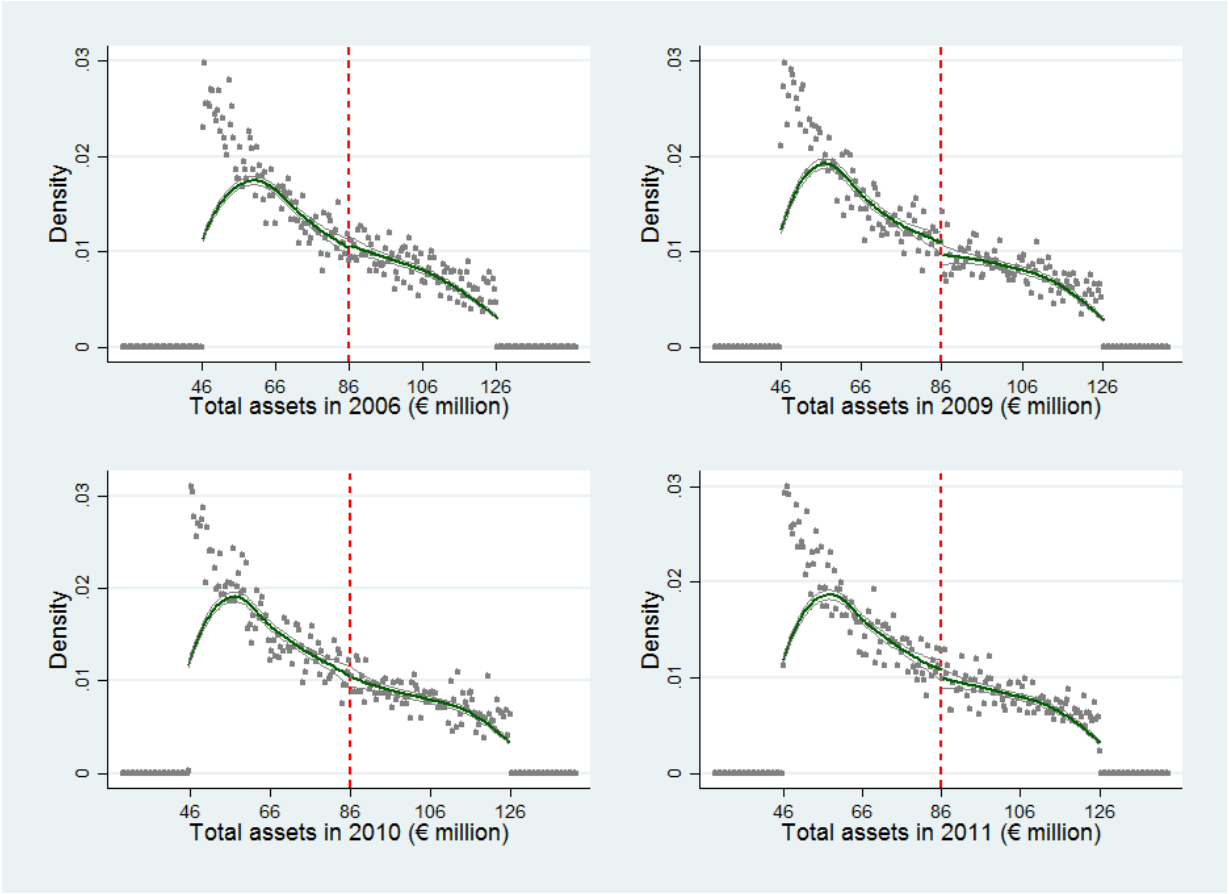
Note: Discontinuity estimate at each placebo threshold is estimated using the baseline first-stage R&D expenditure regression (OLS Regression Discontinuity Design (RDD) with average R&D expenditure over 2009-11 as the dependent variable).The running variable is total assets in 2007. Baseline sample includes firms with total assets in 2007 €25m above and below the placebo threshold. Controls for first order polynomials of running variable separately for each side of the placebo threshold are included. The dashed lines indicate the 95% confidence interval for the discontinuity estimates.

Figure A2. Discontinuities in average number of patents over 2009-11 at "pseudo" SME asset thresholds



Note: Discontinuity estimate at each placebo threshold is estimated using the baseline reduced-form R&D expenditure regression (OLS estimates based on the RD design with average number of patents over 2009-11 as the dependent variable).The running variable is total assets in 2007. Baseline sample includes firms with total assets in 2007 €25m above and below the placebo threshold. Controls for first order polynomials of running variable separately for each side of the placebo threshold are included. The dashed lines indicate the 95% confidence interval for the discontinuity estimates.

Figure A3. McCrary tests for no manipulation at the SME asset threshold, year-by-year

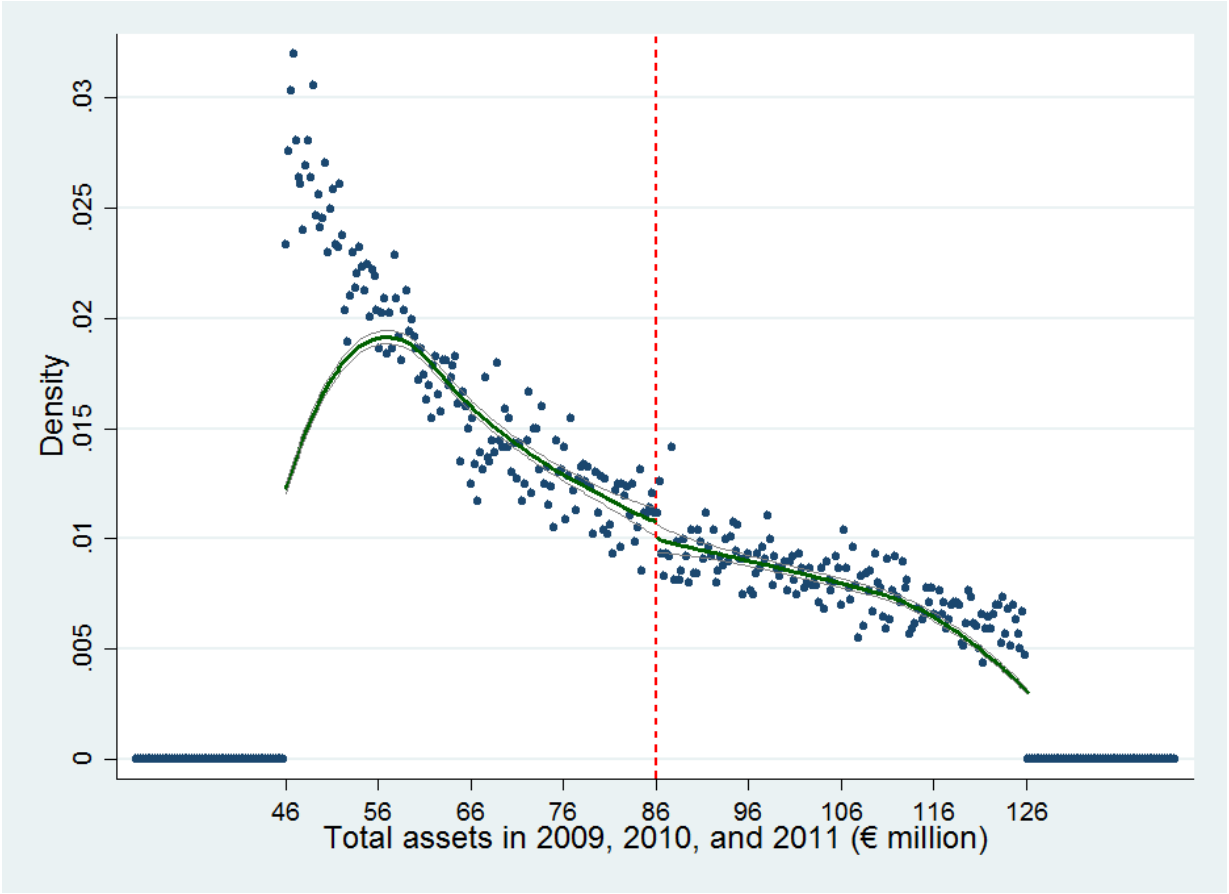


Note: McCrary tests for discontinuity in distribution density of total assets at the SME asset threshold of €86m, year-by-year for 2006, 2009, 2010, and 2011. Each sample includes firms with total assets in [€46m, €126m] in the respective year. The discontinuity estimate (log difference in density height at the SME threshold) (standard error) in 2006 is 0.029 (0.065), in 2009 is -0.125 (0.078), in 2010 is -0.006 (0.077), and in 2011 is -0.086 (0.075).

Figure A4. McCrary test for no manipulation at the SME asset threshold before the policy change

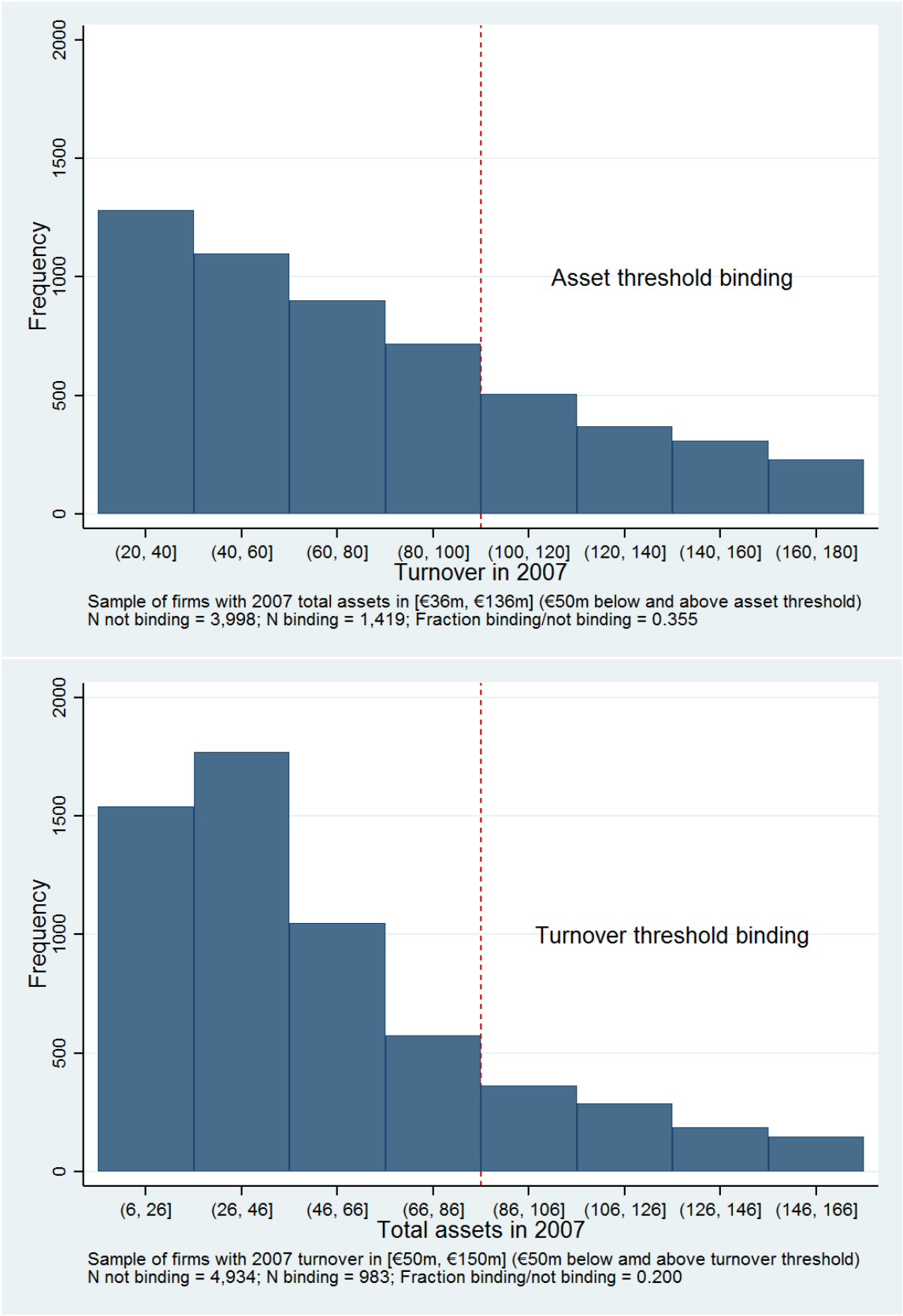
Note: McCrary test for discontinuity in distribution density of total assets at the SME asset threshold of €86m before the policy change, pooling together total assets in 2006 and 2007. Sample includes firms with total assets in [€46m, €126m] in each of the year. The discontinuity estimate (log difference in density height at the SME threshold) is 0.013, with standard error of 0.056.

Figure A5. McCrary test for no manipulation at the SME asset threshold after the policy change



Note: McCrary test for discontinuity in distribution density of total assets at the SME asset threshold of €86m after the policy change, pooling together total assets in 2009, 2010, and 2011. Sample includes firms with total assets in [€46m, €126m] in each of the year. The discontinuity estimate (log difference in density height at the SME threshold) is -0.072, with standard error of 0.045.

Figure A6. Number of firms with binding and not-binding asset and revenue thresholds



Note: Asset threshold is not binding for firms with 2007 sales in (€20m, €100m] and binding for firms with 2007 sales in (€100m, €180m]. Sales threshold is not binding for firms with 2007 total assets in (€6m, €86m] and binding for firms with 2007 total assets in (€86m, €166m].