Online Appendix: Tax Credits and Small Firm R&D Spending

By Ajay Agrawal and Carlos Rosell and Timothy Simcoe

Draft: September 24, 2019

This online appendix provides supplementary analysis and results to accompany the article "Tax Credits and Small Firm R&D Spending" (forthcoming in American Economic Journal: Policy). Please refer to the paper for a more detailed explanation of the data and analysis.

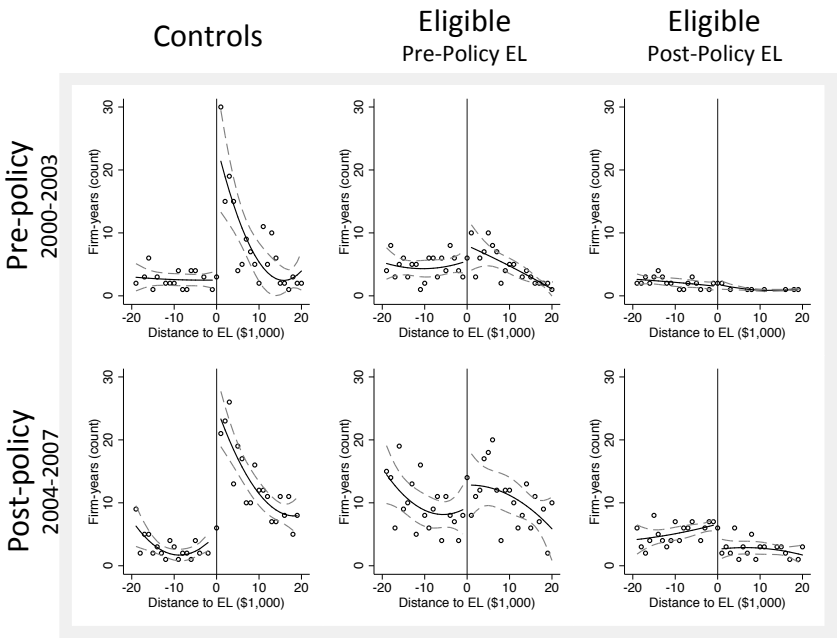
Appendix A

This appendix illustrates that there is an increase in the probability density of firms investing in R&D near the expenditure limit, and proposes a simple model along the lines of Garicano et al. (2016) to rationalize this behavior. We begin by showing that when firms cross the expenditure limit threshold, they respond to the increased marginal cost of R&D.

Figure A-1 is divided into six panels, with the top row corresponding to data from the pre-policy time period, and the bottom row using data from after the policy change. Each graph shows a count (or probability density) of firm-year observations conditional on distance from the expenditure limit, which is indicated by a vertical line. The dots are actual frequencies (i.e., the number of firm-year observations where R&D expenditure is at a certain level relative to the threshold). The lines correspond to fitted values and a 95 percent confidence interval from a quadratic model with a break-point at the expenditure limit.[1]

[1]To produce each panel in Figure A-1, we first created a variable X_{it}, equal to firm i's *Total R&D* in year t minus the relevant expenditure limit EL_{it}. Next, using observations where $|X_{it}| <$ \$1 million,

FIGURE A-1. R&D INVESTMENT NEAR THE EXPENDITURE LIMIT



The first column in Figure A-1 shows that for "Control" firms with lagged taxable income below $200 thousand, there is a large jump in the probability density right at the expenditure limit. This jump in probability mass above the EL threshold suggests that the size of the marginal R&D investment falls when the after-tax marginal cost of R&D increases, making firm-year observations less "spread out" along the X-axis.

The middle column examines firm-year observations in the "Treatment" group relative to the pre-policy expenditure limit. Comparing these graphs to the left-hand column, it appears that firms in the "Treatment" sample are less sensitive to the expenditure limit. However, there is a notable increase in the frequency (density) of observations as we move from the pre to the post-policy time-period

we counted the number of firm-years where X_{it} fell into each of a series of 80 "bins" with a bandwidth of $25,000. Formally, letting $k = -39 \dots 40$ index the bins, we created variables $Y_k = \sum_{i,t} 1[25,000 * (k-1) < X_{it} \leq 25,000 * k)]$ and $X_k = 25k$. We then created scatter plots of the 80 values of (Y, X), along with fitted values and 95% confidence intervals from the regression: $Y = \alpha + \beta_1 X + \beta_2 X^2 + 1[X > 0]\{\alpha_2 + \gamma_1 X + \gamma_2 X^2\} + \varepsilon_k$.

(i.e. from the top to the bottom row) in the middle column. This shows that firms are more willing to spend at or above the pre-policy expenditure limit *after that limit has been increased*. This is the source of variation used to identify the main results in the body of the paper.

Finally, the right-hand column in Figure A-1 shows the probability density in a neighborhood of the post-policy expenditure limit. In the top-right cell, we can see that there are very few firm-year observations from the pre-policy period that spend enough on R&D to reach the post-policy threshold. However, in the bottom-right cell, we can observe not only more observations, but also a discontinuous drop in the density at the expenditure limit. We interpret this bunching as evidence that firms are aware of the change in the SRED policy, and are adjusting their spending to take advantage of the higher threshold.

Because the bunching of Control observations just above the expenditure limit in the left-most column of Figure A-1 may seem counter-intuitive, we now provide a simple model to illustrate the the incentives behind this behavior. The model rests on two assumptions: (1) firms differ in their marginal productivity of R&D, and (2) there is a sharp increase in the marginal cost of R&D at the expenditure limit.

Suppose that a firm investing x in R&D receives gross benefits $B(x;\eta) = \eta x^{\theta}$, where $\theta < 1$ and η is a random parameter with cumulative distribution F (Assumption 1). Further, suppose that the marginal cost of R&D is c^L up to some expenditure limit EL, and c^H thereafter (Assumption 2), so total costs are $C(x) = c^L \min\{x, EL\} + c^H \max\{0, x - EL\}$. The first-order condition for R&D investment then implies that:

$$x^*(\eta) = \begin{cases} \left[\frac{\theta\eta}{c}\right]^{1/(1-\theta)} & \text{if } \eta \leq \underline{\eta} \text{ or } \eta \geq \overline{\eta} \\ EL & \text{if } \underline{\eta} < \eta < \overline{\eta}, \end{cases} \tag{A-1}$$

where $\underline{\eta} = c^L \frac{EL^{(1-\theta)}}{\theta}$, and $\overline{\eta} = c^H \frac{EL^{(1-\theta)}}{\theta}$. Thus, Assumptions 1 and 2 suffice to

generate a mass point in the distribution of x at the expenditure limit, since there is an atom of types $[\underline{\eta}, \overline{\eta}]$ that spend exactly $x^* = EL$.

Now let $g(x)$ denote the density of x. Applying the chain rule, we have $g(x) = F'(\eta(x))\eta'(x)$, where $\eta(x)$ is the inverse of the optimal R&D investment implied by the first-order condition (A-1). Taking the limits of $g(x)$ from above and below as x approaches EL, we have

$$\text{(A-2)} \qquad \lim_{x \uparrow EL} = F'(\underline{\eta}) c^L \frac{(1-\theta)}{\theta EL^\theta} \text{ and } \lim_{x \downarrow EL} = F'(\overline{\eta}) c^H \frac{(1-\theta)}{\theta EL^\theta}$$

Thus, a graph of $g(x)$ will have a discontinuous increase at $x = EL$ if and only if $c^H F'(\overline{\eta}) > c^L F'(\underline{\eta})$. That is what we observe in the lefthand column of Figure A-1, for firm-years in the Control sample (i.e. with lagged taxable income below $200 thousand). Intuitively, we see bunching above the cutoff because the increase in marginal cost leads more firms to "drop out" for each additional dollar of R&D investment, and this produces an increase in the density $g(x)$ at the expenditure limit.

Although Figure A-1 provides strong evidence that firms respond to the R&D tax credit, we do *not* use this variation in our empirical analysis for two reasons. First, the marginal cost of R&D is endogenous – our main outcome variable appears on the x-axis in Figure A-1.[2] And second, only about two percent of the firm-year observations in our data set actually cross the expenditure limit threshold.

[2]This simultaneity also prevents us from using a regression discontinuity design to estimate the impact of tax credits at the expenditure limit.

Appendix B

TABLE B-1—CANADIAN-CONTROLLED PRIVATE CORPORATION MARGINAL TAX RATES

	2000	2001	2002	2003	2004	2005	2006	2007
Maximum small business limit ($thous.)	$200	$200	$200	$225	$250	$300	$300	$400
Tax rate up to reduced business limit†	13.12	13.12	13.12	13.12	13.12	13.12	13.12	13.12
Tax rate from reduced business limit to $300K	29.12	22.12	22.12	22.12	22.12	22.12	22.12	22.12
Tax rate above $300K or small-business deduction threshold	29.12	28.12	26.15	24.12	22.12	22.12	22.12	22.12

†The reduced business limit varies between $0 and the maximum small business deduction threshold depending on the firm's size as determined by taxable capital employed in Canada.

TABLE B-2—IMPACTS OF SRED FOR BALANCED PANEL

Specification: Poisson QML Regression Unit of Analysis: Firm-Year						
Outcome Variable	Total R&D (1)	Total R&D (2)	Total R&D (3)	R&D Wages (4)	R&D Contracts (5)	Non-R&D Investment (6)
Eligible X Post policy	0.15	0.18	0.14	0.12	0.23	0.11
	(0.06)	(0.05)	(0.05)	(0.05)	(0.10)	(0.10)
Eligible	0.11	0.07	0.03	0.01	0.09	0.12
	(0.04)	(0.04)	(0.04)	(0.04)	(0.09)	(0.08)
Post policy	0.15					
	(0.03)					
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes	Yes
Psuedo-R2	0.75	0.75	0.80	0.84	0.68	0.58
Observations	35,101	35,101	35,101	27,424	26,020	34,129
Number of firms	4,495	4,495	4,495	3,515	3,326	4,364
Mean of outcome variable	73,018	73,018	73,018	64,468	14,448	87,152

Notes: Robust standard errors (clustered by firm) in parentheses. All models are estimated using a balanced panel of N=35,101 firm-years; changes in sample size are due to omission of any firm with all-zero outcomes. The mean value of the outcome variable is calculated for all firm-years used in these estimations.

Table B-3—Dropping Observations Above Expenditure Limit

Specification: Poisson QML Regression Unit of Analysis: Firm-Year				
Outcome Variable	Total R&D (1)	Total R&D (2)	R&D Wages (3)	R&D Contracts (4)
Eligible X Post policy	0.13	0.12	0.07	0.29
	(0.05)	(0.04)	(0.04)	(0.10)
Eligible	0.11	0.04	0.05	0.06
	(0.04)	(0.03)	(0.03)	(0.08)
Firm FE	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes
Psuedo-R2	0.72	0.82	0.84	0.65
Observations	47,579	47,579	37,711	35,299
Number of firms	7,191	7,191	5,755	5,326
Mean of outcome variable	72,360	72,360	61,050	16,993

Notes: Robust standard errors (clustered by firm) in parentheses. The mean value of the outcome variable is calculated for all firm-years used in these estimations.

Table B-4—Adjustment Cost Estimates for Balanced Panel

Specification: Poisson QML Regression
Unit of Analysis: Firm-Year

Sample	All Firm-Years			Non-NAICS 541 Firm-Years		
R&D Outcome Variable	Total	Wages	Contracts	Total	Wages	Contracts
Eligible X Policy X Zero-tax	0.15	0.19	0.26			
	(0.09)	(0.09)	(0.23)			
Policy X Zero-tax	-0.23	-0.17	-0.26			
	(0.04)	(0.04)	(0.09)			
Eligible X Zero-tax	-0.08	-0.11	-0.13			
	(0.08)	(0.07)	(0.17)			
Zero-tax-liability	0.15	0.13	0.10			
	(0.03)	(0.03)	(0.08)			
Eligible X Policy X Capital				0.28	0.22	0.38
				(0.12)	(0.10)	(0.26)
Policy X Capital				-0.27	-0.16	-0.34
				(0.07)	(0.05)	(0.13)
Eligible X Capital				-0.16	-0.16	-0.09
				(0.09)	(0.09)	(0.20)
Eligible X Policy	0.02	0.01	0.06	-0.04	-0.04	0.03
	(0.05)	(0.05)	(0.12)	(0.05)	(0.06)	(0.17)
Eligible	0.08	0.07	0.16	0.11	0.11	0.09
	(0.04)	(0.04)	(0.11)	(0.04)	(0.05)	(0.13)
Additional controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Psuedo-R2	0.80	0.84	0.68	0.81	0.84	0.67
Observations	35,101	27,424	26,020	26,251	19,273	19,945
Total Firms	4,495	3,515	3,326	3,350	2,463	2,542
Mean of outcome	73,018	64,468	14,448	61,821	56,343	10,600

Notes: Robust standard errors (clustered by firm) in parentheses. All models are estimated using a balanced panel of N=35,101 firm-years; changes in sample size are due to omission of any firm with all-zero outcomes. The mean value of the outcome variable is calculated for all firm-years used in these estimations.

TABLE B-5—SAMPLES EXCLUDING YEARS AROUND SRED POLICY-CHANGE

Specification: Poisson QML Regression
Unit of Analysis: Firm-Year
Outcome Variable: Total R&D

Sample	Full Sample (1)	Drop 2003 (2)	Drop 2004 (3)	Drop 03-04 (4)	Drop 03-05 (5)
Eligible X Post policy	0.18	0.19	0.22	0.22	0.22
	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)
Eligible	0.00	-0.02	0.00	0.00	0.03
	(0.03)	(0.05)	(0.03)	(0.05)	(0.05)
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Observations	48,638	40,185	41,906	32,487	25,437
Number of firms	7,239	6,850	7,170	6,486	6,116

Notes: Robust standard errors (clustered by firm) in parentheses. Each estimate corresponds to the model in Column (3) of Table (2), estimated on samples that exclude different combinations of pre and post-policy years.

TABLE B-6—OLS SPECIFICATION FOR SRED POLICY IMPACT

Specification: Ordinary Least Squares
Outcome: log(max{Total R&D, X})

Outcome: X =	Missing (1)	$10,000 (2)	$25,000 (3)	$50,000 (4)
Eligible X Post policy	0.11	0.18	0.12	0.09
	(0.03)	(0.03)	(0.02)	(0.02)
Eligible	0.03	0.06	0.06	0.05
	(0.03)	(0.03)	(0.02)	(0.02)
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
R-squared	0.93	0.72	0.75	0.77
Observations	28,713	48,638	48,638	48,638
Number of firms	7,239	7,239	7,239	7,239

Notes: Robust standard errors (clustered by firm) in parentheses. All models are estimated using an unbalanced panel of all available firm-years. Model (1) drops observations with no reported R&D expenditures.

TABLE B-7—ALTERNATIVE MEASURE OF ZERO-TAX-LIABILITY (NO CURRENT REVENUE)

Specification: Poisson QML Regression
Unit of Analysis: Firm-Year

R&D Outcome Variable	Total (1)	Wages (2)	Contracts (3)
Eligible X Policy X Zero-tax	0.28	0.19	0.51
	(0.10)	(0.10)	(0.24)
Policy X Zero-tax	-0.28	-0.18	-0.44
	(0.04)	(0.04)	(0.08)
Eligible X Zero-tax	-0.15	-0.13	-0.24
	(0.08)	(0.08)	(0.17)
Eligible X Policy	0.02	0.03	0.07
	(0.04)	(0.04)	(0.11)
Eligible	0.08	0.06	0.15
	(0.03)	(0.03)	(0.09)
Zero-tax-liability	0.11	0.07	0.11
	(0.03)	(0.03)	(0.06)
Additional controls	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes
Psuedo-R2	0.81	0.83	0.67
Observations	48,638	38,748	36,235
Total Firms	7,239	5,806	5,378
Mean of outcome	82,887	69,310	18,895

Notes: Robust standard errors (clustered by firm) in parentheses. All models are estimated using an unbalanced panel of all available firm-years; changes in sample size occur when firms with all-zero outcomes are dropped from the conditional fixed-effects specification. The mean value of the outcome variable is calculated for all firm-years used in the estimation.

TABLE B-8—MEANS AND SAMPLE SIZES FOR TABLE 5

	All Firms	Taxes Owed	Zero Tax Liability
Pre-2004 Mean Outcome			
Total R&D	161,291	124,785	322,983
R&D Wages	106,075	87,294	189,734
R&D Contracts	26,403	20,407	53,003
Post-2003 Eligible Observations			
Observations	2,346	1,975	371

REFERENCES

Garicano, L., C. Lelarge, and J. Van Reenen (2016). Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review forthcoming*.