Firm Heterogeneity and the Aggregate Effects of Innovation Policy

Marek Ignaszak* Daniel Robbins† Petr Sedláček‡

August 6, 2026

Abstract

R&D elasticities, which measure how firms adjust R&D in response to changes in tax incentives, are central to the design of innovation policy. Exploiting firms optimal R&D decisions, we characterize these elasticities using sufficient statistics. Unlike existing methodologies, our approach does not rely on policy variation and recovers firm-specific elasticities rather than only average responses for broad groups of firms. We validate our approach against existing methodologies and apply it to Compustat data. Accounting for differences in R&D elasticities doubles the aggregate benefits of R&D incentives, with young, fast-growing firms delivering the largest increase in aggregate growth per subsidy dollar.

JEL Classification: O31, O38, L1

Keywords: Heterogeneous firms, Economic growth, R&D incentives

*Goethe University Frankfurt, email: ignaszak@uni-frankfurt.de.

†University of New South Wales, email: daniel.robbins@student.unsw.edu.au

‡Corresponding Author; University of New South Wales, Cnr Union Rd & College Rd, UNSW Sydney, Kensington NSW 2033, Australia; tel: +61 2 9385 9903; email: p.sedlacek@unsw.edu.au.

We thank Ufuk Akcigit, Matěj Bajgar, Christian Bayer, Adrien Bilal, Andrea Chiavari, James Cloyne, Gustavo de Souza, Georg Dürnecker, Wouter den Haan, Niklas Engbom, Owen Freestone, Leo Kaas, Greg Kaplan, Marta Kozakiewicz, Danial Lashkari, Omer Majeed, Federico Masera, Luca Mazzone, Vincent Sterk, Emily Swift and seminar participants at the 2nd Macro Development Workshop, Australian Conference of Economists, Australian National University, Australian Treasury, Australian Department of Industry, Science and Resources, Bristol Macroeconomic Workshop, Dutch National Bank, Goethe University Frankfurt, IAER Workshop at Dongbei University, National University of Singapore, Nanyang Technological University, New York University Abu Dhabi, OECD, Peking University HSBC Business School, Productivity Research Conference in Manchester, Singapore Management University, Society of Economic Dynamics Annual Meeting in Athens, University of Bonn, University of New South Wales and the Workshop on Firm Dynamics and Labor in Copenhagen for useful comments. Sedláček gratefully acknowledges the financial support of the Australian Research Council [grant number FT230100545] and, in early stages of the project, of the European Research Council [grant number 802145]. The results of these studies are based, in part, on data supplied to the ABS under the Taxation Administration Act 1953, A New Tax System (Australian Business Number) Act 1999, Australian Border Force Act 2015, Social Security (Administration) Act 1999, A New Tax System (Family Assistance) (Administration) Act 1999, Paid Parental Leave Act 2010 and/or the Student Assistance Act 1973.

1 Introduction

Innovation is widely viewed as a key driver of productivity growth and sustained improvements in living standards. Motivated partly by the positive externalities generated by innovation, governments around the globe devote substantial public resources to stimulating private research and development (R&D) through a range of incentives (see e.g. Bloom et al., 2019; OECD, 2023). Designing and evaluating such R&D policies, however, requires understanding firms' "R&D elasticities" – how businesses adjust their innovative activity in response to changes in R&D incentives.

A large empirical literature estimates R&D elasticities by exploiting policy variation across time, regions or eligibility thresholds (see e.g. Hall and van Reenen, 2000; Rao and Simcoe, 2026, for a discussion). However, these approaches typically recover only *average* responses for broad groups of firms around specific policy changes. Instead, we develop a novel methodology for measuring R&D elasticities of *individual firms* without relying on policy variation. To showcase the type of analysis made possible by our approach, we use it to study the aggregate implications of heterogeneity in firm-level R&D elasticities.

Our approach leans on fundamental economic trade-offs faced by firms when deciding on optimal R&D investment – a balance between the associated marginal costs and benefits. Under conditions consistent with a large class of models of endogenous growth (see e.g. Klette and Kortum, 2004; Lentz and Mortensen, 2008; Acemoglu and Cao, 2015), we show analytically that firms' R&D elasticities can be summarized by sufficient statistics – simple firm-specific moments measurable with readily available micro-data.

Importantly, we validate our methodology by showing that the proposed sufficient statistics provide very similar estimates to those obtained from established approaches utilizing quasi-experimental variation in policies (see e.g. Appelt et al., 2025). In contrast to existing approaches, however, our methodology has two key advantages: we can measure R&D elasticities (i) at the firm-level and (ii) without requiring policy variation.

To highlight these benefits, we apply our approach to Compustat data. Three results stand out. First, firm heterogeneity matters. The inability to measure R&D elasticities at the firm-level—the typical case in the existing literature—understates the positive impact such policies have on aggregate growth by one half. This is because R&D elasticities are not the sole driver of the aggregate impact. Instead, what matters is the joint distribution of firms' R&D elasticities with their sizes and growth rates. In the data, firms which are responsive to R&D policies also tend to grow faster, strengthening their aggregate impact. Second, we show that young, fast-growing businesses deliver the largest relative increase in aggregate growth per dollar spent on the R&D incentive. Finally, we document that our results do not change even when accounting for knowledge spillovers across businesses, dynamics, or borrowing constraints.

More specifically, we begin our analysis by focusing on firms' optimal R&D investment decisions. We assume that, in order to grow, firms invest into R&D by optimally balancing associated marginal costs and benefits. While the former are governed by firms' cost structures, the latter are driven by changes in firms' profits (firm value) stemming from successful innovation. In this environment, we consider a permanent change in the (tax) price of R&D – representing changes in government innovation incentives – and derive the associated R&D elasticity.

We show analytically that, when firms' values are proportional to sales – a condition satisfied by a large class of endogenous growth models – firms' R&D elasticities are proportional to the ratio of their present discounted value of R&D costs to firm value. Moreover, in a special case, firms' R&D elasticities become proportional to "just" firms' prevailing R&D cost to profit ratios.[1] Intuitively, as R&D becomes cheaper due to more generous innovation incentives, firms optimally increase their R&D investment. This has two opposing effects. On the one hand, greater R&D investment is costly. On the other hand, it raises expected growth and, therefore, future profits and firm value. The ratio of these two sides of the balancing act pins down firms' R&D elasticities.

A crucial step in our analysis is to validate the proposed sufficient statistics. We do so in two ways. First, using Australian firm-level data (BLADE) we leverage time variation in R&D incentives around a large policy change in 2012 and estimate firms' R&D elasticities with a conventional event study design. The results show that in response to the 2012 increase in R&D tax incentives, (treated) firms invested considerably more into R&D. More importantly for the purpose of this paper, our sufficient statistics – computed for the same set of treated firms – lie comfortably within the estimated confidence bands.

Second, we use the Australian micro-data to more *directly* test our sufficient statistics. In particular, employing a conventional difference-in-differences design we estimate heterogeneous treatment effects of the 2012 policy change with respect to firms' R&D to profit ratios – our sufficient statistic. The theory suggests that the estimated heterogeneity coefficient should be equal to one, i.e. firms' R&D elasticities move one-for-one with our proposed sufficient statistics. Indeed, in the data we cannot reject the null hypothesis that this is the case.[2]

To showcase the type of analysis our methodology makes possible, we apply it to Compustat data and study the aggregate implications of firm-level heterogeneity in R&D elasticities. As a first step, we develop an analytical accounting framework that allows us to understand how firm-level R&D elasticities shape aggregate growth and total government

[1] Examples of such models include, e.g., Klette and Kortum (2004); Luttmer (2007); Lentz and Mortensen (2008); Luttmer (2010); Mukoyama and Osotimehin (2019). The special case is a result of so called "perfect scaling" whereby sales and costs grow one-for-one with firm size (see Akcigit and Kerr, 2018, for a discussion). Appendix D.1 explicitly shows that in our data firms' sales to profit and R&D to profit ratios are approximately constant over their life-cycles.

[2] In addition, we also show that our sufficient statistics deliver very similar results to estimated R&D elasticities in Orbis and Compustat data (see Hall, 1993; Appelt et al., 2025). This is true both for all firms on average, but also across the firm size distribution.

spending on R&D incentives. Intuitively, individual firms can only leave a mark on aggregate growth if they are sufficiently large, fast-growing, and responsive to R&D policy changes. At the same time, changing an R&D policy will be particularly expensive if firms are very responsive to it, or if targeted businesses already receive a substantial portion of government resources. Three key messages emerge.

First, firm heterogeneity matters. In the data, firms characterized by higher R&D elasticities tend to grow faster. This covariance boosts the aggregate benefits of R&D support. Quantitatively, ignoring firm-level differences in R&D elasticities understates the aggregate benefits of R&D incentives by one half. Therefore, the aggregate impact of firm-level R&D incentives is potentially much larger than previously thought.

Second, we turn our attention to various groups of firms at which the government may wish to target its R&D policy. Our results suggest that, out of the considered groups, young and fast-growing firms offer the highest "bang for the buck" – an increase in aggregate growth per dollar spent on the R&D incentive. While small firms display a similarly high bang for the buck, this is largely driven by the fact that the majority of small businesses are young. In fact, small-old firms have a lower bang for the buck than a random sample of businesses. These results align well with existing evidence on the economic prowess of young firms (see, e.g., Haltiwanger et al., 2013, 2016). Our results provide new evidence that young businesses are also key propagators of R&D policies.

In the last step of our analysis, we show that accounting for knowledge spillovers, dynamics, or borrowing constraints does not change our results. To incorporate knowledge spillovers, we extend our analytical framework to allow firm growth to be partly driven by R&D investment of other businesses. Empirically, we build on Bloom et al. (2013) and measure the technological proximity of firms using citation-weighted patents and their technological classification. While accounting for spillovers favors larger, older and more R&D-intensive businesses, young and fast growing firms still deliver the greatest relative benefits per dollar spent. Similarly, our results are robust to accounting for dynamics. Specifically, the data shows that young firms continue to deliver the strongest relative benefits per dollar spent even a decade into the future. Finally, in Appendix B.4 we extend our framework with borrowing constraints and show that while more generous innovation subsidies also alleviate credit constraints, this effect is dwarfed by firms' primary interest in taking advantage of cheaper R&D costs.[3]

Related Literature. Our paper is related to several strands of the literature. First, it connects to empirical studies utilizing policy variation to estimate how firms and innovators respond to (tax) policy changes (see, e.g., Hall, 1993; Bloom et al., 2002; Rao, 2016; Guceri and Liu, 2019; Agrawal et al., 2020; Chen et al., 2021; Akcigit et al.,

[3]Note also that Ottonello and Winberry (2024) estimate that the majority of innovation in the U.S. is performed by unconstrained firms.

2022a; Appelt et al., 2025). While these studies differ in identification strategies, they all estimate *average* responses around specific policy changes. A notable exception is de Souza (2025) who, using a difference-in-differences design on Brazilian micro-data, estimates a distribution of treatment effects as a function of firm characteristics. In contrast, our approach measures R&D elasticities at the *firm level* through sufficient statistics without tying their heterogeneity to other observables and without relying on policy variation or thresholds.

Therefore, our paper belongs to the sufficient statistics tradition (see e.g. Chetty, 2009), which derives policy-relevant elasticities from a small set of estimable moments while remaining agnostic about the full structure of the underlying model. We are closest to Atkeson and Burstein (2019), who derive aggregate sufficient statistics for the dynamic response of aggregate growth to innovation policies. While their analysis focuses on a uniform (policy-induced) increase in firms' innovation, we show that heterogeneity in firms' responses to policy changes is quantitatively important for the aggregate impact. As such, we take a step toward addressing the challenge highlighted by Atkeson and Burstein (2019) of developing firm-level metrics for evaluating where innovation spending should be directed.[4]

Our approach builds on a long tradition of endogenous growth models and studies highlighting the importance of firm heterogeneity for aggregate outcomes (see, e.g., Aghion and Howitt, 1992; Hopenhayn and Rogerson, 1993; Klette and Kortum, 2004; Lentz and Mortensen, 2008; Akcigit and Kerr, 2018; Sterk ⓡ al., 2021) from which we borrow "only" the trade-offs inherent to firms' optimal R&D decisions to derive our sufficient statistics. In contrast to recent model-based evaluations of R&D policies (see e.g. Acemoglu et al., 2018; Akcigit et al., 2022b), our approach can account for the full extent of observed firm-level heterogeneity without the need for simplifying parameterizations.

Paper Organization. While the organization of our paper is designed to improve readability, we caution the reader that it is slightly unorthodox as it switches between theory and data twice. In particular, Section 2 presents all the results pertaining to our sufficient statistics approach – it develops the analytical framework, maps it to the data and validates it against established methodologies. Then, Section 3 studies the aggregate implications of firm heterogeneity in R&D elasticities – it presents the analytical accounting framework, maps it to the data and presents empirical results. Section 4 provides a discussion of potential extensions and the last section concludes.

[4] Our aggregate object of interest – the impact of R&D policy on aggregate growth per public dollar spent – is in the spirit of the marginal value of public funds of Hendren and Sprung-Keyser (2020) and the fiscal cost of innovation subsidies in Atkeson and Burstein (2019). We discuss how our measure connects to welfare in Section 4.

2 Firm-Level Innovation Elasticities

We begin by theoretically analyzing firm-level R&D elasticities – measures of how individual firms adjust their R&D investment in response to a change in (tax) incentives. The key result from this section is that firms' R&D elasticities can be described by sufficient statistics – simple moments measurable in the data that do not require us to specify an entire structural model.[5]

In what follows, we first present the theory and derive our sufficient statistics. Next, we explain how the derived sufficient statistics can be measured using readily-available firm-level datasets. Finally, at the end of this section we validate our approach against estimates of average R&D elasticities using established methodologies (see e.g. Dechezlepretre et al., 2023; Appelt et al., 2025). The next section showcases the advantages of our approach by analyzing the aggregate implications of firm-level heterogeneity in innovation elasticities.

2.1 Theory

Our methodology rests on key economic trade-offs firms face when deciding on R&D investment: the balance of associated marginal costs and benefits. We now formalize these trade-offs.

Firm-Level Innovation. We consider an environment in which heterogeneous firms, indexed by i, optimally invest into R&D in order to grow. In particular, firm growth, g_i, is innovation-driven and assumed proportional to firms' innovation rates, x_i, such that $g_i = \lambda_i x_i$ with $\lambda_i > 0$. The associated R&D investment is denoted by $s_i(x_i) = P_s \widetilde{s}_i(x_i)$ where $\widetilde{s}_i(x_i)$ is the "quantity" and $P_s \leq 1$ is the (tax) "price" of R&D. While the latter is controlled by the government, the former is the firm's choice and assumed to be convex in the innovation rate:

$$\frac{\partial s_i(x_i)}{\partial x_i} \frac{x_i}{s_i(x_i)} = \psi_i, \tag{1}$$

where $\psi_i > 1$ is the R&D "cost elasticity." In this environment, firms optimally choose innovation rates in order to maximize their value, $v_{i,t}$:

$$v_{i,t} = \max_x \left[\pi_{i,t} + \beta_i v_{i,t+1} \right], \tag{2}$$

where $\pi_{i,t} = y_{i,t} - w_{i,t} - s_{i,t}$ are profits and where we explicitly separate R&D expenditures, $s_{i,t}$, from sales, $y_{i,t}$, and all other operating costs, $w_{i,t}$. In the above, $\beta_i \in (0, 1)$ is a

[5]We focus on the intensive margin of R&D as Dechezlepretre et al. (2023) find that the extensive margin does not significantly respond to R&D incentives. Appendix B.1 lays out how our theory maps into a fully-fledged general equilibrium endogenous growth model and Appendix B.2 explicitly derives that our sufficient statistics hold in canonical growth frameworks, such as Klette and Kortum (2004).

discount factor, incorporating aggregate growth and potentially also firm exit.[6]

To study firms' optimal innovation decisions, we draw on a large class of endogenous growth models in which firm value is proportional to sales (see e.g. Klette and Kortum, 2004; Lentz and Mortensen, 2008; Acemoglu and Cao, 2015; Mukoyama and Osotimehin, 2019). In this setting, optimal R&D decisions can be described by the following condition:

$$\overbrace{\psi_i \frac{s_{i,t}}{x_{i,t}}}^{\text{marginal cost}} = \overbrace{\beta_i \frac{\lambda_i}{1+g_{i,t}} v_{i,t+1}}^{\text{marginal benefit}} . \tag{3}$$

Intuitively, on the one hand marginal costs are governed by the firm's R&D cost function. On the other hand, marginal benefits reflect the additional increase in firm value stemming from successful innovations, where $\lambda_i/(1+g_{i,t})$ is the semi-elasticity of a firm's gross growth with respect to its R&D.[7]

R&D Elasticities. We are now ready to derive our key *firm-level* objects of interest. In doing so, we explicitly make a distinction between two separate, but closely related, elasticities. First, we define the "innovation elasticity" which measures how firms adjust their innovation rates in response to a change in the tax price of R&D:

$$\epsilon_i = \frac{\partial x_i}{\partial P_s}\frac{P_s}{x_i} = \frac{\partial g_i}{\partial P_s}\frac{P_s}{g_i} \leq 0. \tag{4}$$

Next, we specify how innovation elasticities relate to "R&D elasticities" which measure how firms adjust their R&D expenditure in response to a change in the tax price of R&D. R&D elasticities are the typical objects of interest in empirical studies and in our framework they are given by:[8]

$$\epsilon_i^s = \frac{\partial \widetilde{s}_i}{\partial P_s}\frac{P_s}{\widetilde{s}_i} = \psi_i \epsilon_i \leq 0. \tag{5}$$

The following proposition uses equation (3) to derive firms' innovation elasticities.

Proposition 1 (Firm-level innovation elasticities). *In the environment described above, firm-level innovation elasticities are:*

[6]While β_i may reflect firm-specific *risk* of shutting down, our theory is derived under the assumption that $v_i \geq 0$ as firms with $v_i < 0$ are typically assumed to exit in full models of endogenous growth.

[7]For tractability, we describe our theory under perfect foresight. Appendix B.3 provides an extension with stochastic profits, showing that our key results carry over. Section 4 and Appendix B.4 extend our framework to consider possible credit constraints, forcing a wedge between the marginal costs and benefits in (3). Appendix D.1 shows that in our data, profits (firm values) are approximately proportional to sales.

[8]Here $\widetilde{s}_i$ represents pre-subsidy R&D expenditures, which is what both tax records and typical micro-data (e.g. Compustat's `xrd`) report. Our ϵ_i^s is therefore directly comparable to the R&D elasticities estimated in the literature. Note also that at the evaluation point $P_s = 1$, $s_{i,t} = \widetilde{s}_{i,t}$. Finally, when deriving innovation elasticities, we do so for a theoretical, marginal change in R&D tax incentives. Details of real-life policies will become relevant at the validation stage in Section 2.3.

(i) In the general "present discounted value (PDV) case" given by:

$$\epsilon_i = -\frac{s^c_{i,t+1}}{v_{i,t+1}} \frac{1}{\psi_i - 1}, \tag{6}$$

where $s^c_{i,t+1} = \mathbb{E}_t \sum_{j=1}^{\infty} \beta_i^{j-1} s_{i,t+j}$ *is the present discounted value of firms' future R&D costs.*
(ii) In the "special case," where $s_{i,t}/\pi_{i,t} = s_i/\pi_i$ *for all* t*, given by:*

$$\epsilon_i = -\frac{s_i}{\pi_i} \frac{1}{\psi_i - 1}. \tag{7}$$

The first part of Proposition 1 shows that firm-level innovation elasticities are proportional to the share of the present discounted value of firms' future R&D costs in firm value. Intuitively, as R&D becomes cheaper due to more generous government incentives, firms optimally increase their investment into innovation with two opposing effects. On the one hand, doing more R&D comes at higher costs. On the other hand, it increases expected growth and, therefore, future profits (firm value). The ratio of these two sides of the balancing act pins down firms' innovation elasticities.

The second part of Proposition 1 considers a special case in which firms' R&D to profit ratios are constant over time (though potentially heterogeneous across firms). Aside from its empirical convenience, this special case is of interest because it is a feature of the wide range of endogenous growth models discussed above. In this case, firms' innovation elasticities are simply proportional to their R&D-to-profit ratios.

A key implication of Proposition 1 is that firms' innovation elasticities are measurable using "just" information on their R&D expenditures, s_i, profits (values), π_i (v_i), and an estimate of their cost elasticity, ψ_i. This makes our approach particularly appealing compared to existing methodologies which rely on policy variation (over time, across regions or eligibility thresholds) in order to estimate an *average* innovation elasticity across large groups of (treated) firms.

Relation to Existing Concepts. Before moving on, we reiterate how our elasticities relate to concepts in the literature. In particular, existing studies typically estimate versions of average R&D elasticities, $\bar{\epsilon}^s$. As explained above, our approach allows us to measure R&D elasticities for *individual firms* as $\epsilon^s_i = \psi_i \epsilon_i$.

R&D elasticities are sometimes also referred to as "R&D user cost elasticities" measuring how R&D expenditures respond to changes in the user cost of R&D or the "B-index" (see e.g. Hall et al., 2010; Appelt et al., 2025). The B-index (break-even index) is the after-tax cost of a dollar invested into R&D which, by construction, depends on R&D tax incentives. The user cost of R&D is then typically defined as the B-index multiplied by the sum of the real interest rate and the rate of depreciation (see Rao and Simcoe, 2026, for a discussion).

2.2 Estimation

We now take our theory to the data and provide estimates of innovation elasticities at the firm-level. Towards this end, we first estimate firms' cost elasticities, ψ_i. In what follows, we describe the adopted methodology, (Compustat) data and results.[9]

Methodology. As alluded to above, the first step in our approach is to obtain estimates of firms' cost elasticities, ψ. Towards this end, we follow Hall and Ziedonis (2001) and estimate the R&D cost function, $s_i(x_i)$, where we use patent counts as a proxy for the innovation rate, x_i. Formally, we estimate the following Poisson regression:[10]

$$\mathbb{E}\Big[\text{patents}_{i,t}\Big] = \exp\Big(\alpha_{i,t} + \gamma \log s_{i,t}\Big), \tag{8}$$

where $\mathbb{E}\left[\text{patents}_{i,t}\right]$ is the expected number of patent applications of firm i in period t, and where the patent count is assumed to be Poisson distributed and dependent on fixed effects, $\alpha_{i,t}$, and the firm's R&D expenditures, $s_{i,t}$. In this setting, which we will estimate at different levels of aggregation, the estimated coefficient γ in regression (8) is a measure of the cost elasticity, $\gamma = 1/\psi$.

Next, we measure firms' innovation elasticities, ϵ. Recall from Proposition 1 that there are two versions of firms' innovation elasticities: (i) "PDV-based" – see equation (6) and (ii) "special case" – see equation (7). The latter, special case, is readily measurable with information on cost elasticities, ψ_i, firm-level R&D expenditures, s_i, and profits, π_i. The former, PDV-based elasticities, also require information on firm values, v_i, and associated discount rates, β_i.

Data: R&D, Patents, Profits and Discount Rates. Throughout this subsection, we use Compustat data which has been extensively used in the literature to study firms' innovation (see e.g. Cavenaile et al., 2021; Ignaszak and Sedláček, forthcoming). Our primary sample period is 1980-2019 when R&D coverage is high. During this period, Compustat firms accounted, on average, for 82% of annual aggregate R&D investment and 63% of real GDP.

For the purpose of estimating R&D cost elasticities, we make use of firms' information on R&D expenditures, s_i, represented in Compustat by the `xrd` variable. We merge this data with firm-level information on (eventually granted) patent applications collected by Kogan et al. (2017). Appendix C.1 provides details on both datasets and their linking.

To measure firms' innovation elasticities, we need also information on their profits and – for the PDV-case – also firm values and associated discount rates. In what follows,

[9]Appendices C and D provide, respectively, more information on the datasets and show additional empirical results and a range of robustness checks.

[10]Note that we count only patents that were eventually granted. Estimating regression (8) in growth rates instead delivers similar results.

we define profits as `sales-cogs-xrd`, where `cogs` are all costs of goods sold – a measure of variable costs. Appendix D.10 provides robustness checks with respect to alternative definitions of profits.

As pointed out by e.g. Gormsen and Huber (2025), market value is not necessarily the same as firm value used by businesses in their decision making. Therefore, we measure firms' values as the present discounted value of their (realized) profits, i.e. $v_{i,t} = \sum_{j=0}^{\infty} \beta_i^j \pi_{i,t+j}$. Towards this end, we make use of Gormsen and Huber (2025) who use corporate conference calls to estimate firms' perceived cost of capital and discount rates, β_i, which they then link to Compustat. We then also use these discount rates to compute the net present value of firms' R&D expenditures when computing their PDV-case innovation elasticities – see equation (6).

Data: Sample Selection. Following the convention in the literature, we drop firms in agriculture, finance, insurance, real estate and public utilities. In addition, we drop observations with missing industry classifications, negative sales and observations in which acquisition expenses exceed 10% of revenues. We do the latter in order to control for mergers and acquisitions for which measured growth may only be the result of acquiring a different business unit.

In order to reduce noise in measurement and to allow for time to build in R&D investment and innovation, we average each firm-level variable of interest over non-overlapping 5-year windows. After this time averaging, we pool all observations. Appendix D.11 provides robustness checks with respect to the averaging window.

To deal with outliers, we drop observations for which annual R&D-to-profit ratios exceed $1,000\%$ in absolute value and then winsorize the top 1%. In doing so, the percentiles are computed in a given 2-digit-industry-year cell to avoid leaning on certain industries or periods of time.

Finally, innovation elasticities are derived under the assumption that firm values are positive. Therefore, we drop firm-window observations that exhibit negative measured firm values. The equivalent of this in the "special case" is to drop observations with negative profits. Since this cuts our sample by 21%, Appendix D.13 provides robustness checks for alternative treatment of negative profits.

Results: R&D Cost Elasticities, ψ. Table 1 shows the results when estimating regression (8) at the aggregate level. The estimated coefficients of interest, $\gamma = 1/\psi$, are consistent with the existing literature (see e.g. Hall and Ziedonis, 2001; Bloom et al., 2002). For our application in the next section, we will use estimates of regression (8) at the 2-digit industry level and assign to each firm the corresponding industry-level cost elasticity, see Figure 1.

Table 1: Estimates of the inverse of the R&D cost elasticity ψ

	(I)	(II)	(III)
log R&D expenses, $\gamma = 1/\psi$	0.53	0.63	0.64
	(0.07)	(0.03)	(0.05)
firm fixed effects	✓	✓	✓
time fixed effects	✓		
sector fixed effects	✓		
sector-time fixed effects		✓	✓
additional controls			✓
Observations	11,598	11,512	8,092

Note: Poisson regression estimates of the coefficient γ in regression (8). Standard errors in parentheses are clustered at the firm level. Observations are weighted with log patent citation count. The coefficient corresponds to the elasticity of patents (innovation output) to R&D expenses (innovation inputs). Sample consists of US public firms matched to patent data in Kogan et al. (2017). We restrict the sample to the period 1980 - 2019. To allow for time to build in innovation, we aggregate data into non-overlapping 5 year windows. Each observation corresponds to one firm-window cell. If a firm exists for less than 5 years, we retain the firm in the sample and use all years in which the firm is observed. Additional controls capture financial frictions through liquidity (debt to current assets) and net leverage (debt net of current assets to total assets), as well as dynamics through lagged R&D expenses.

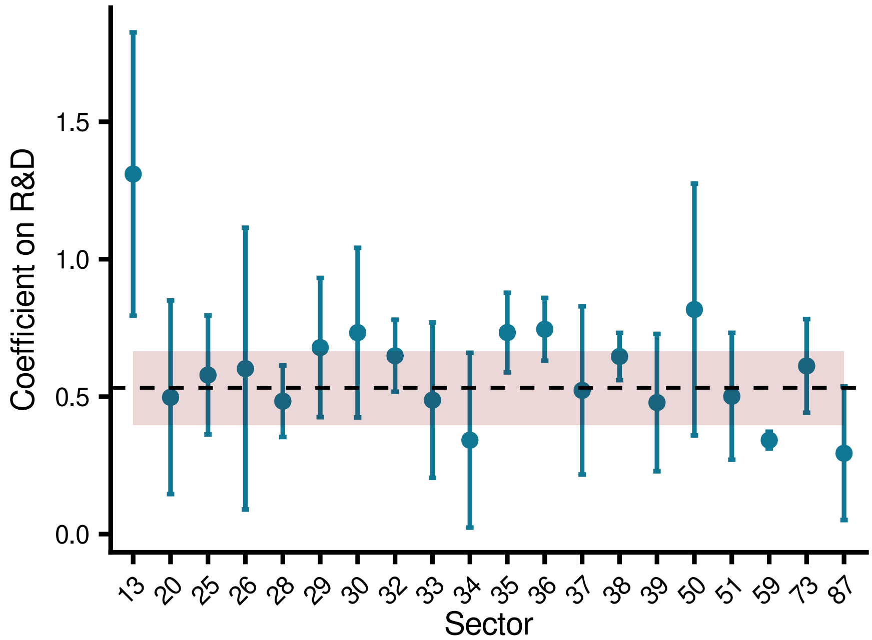
Results: Innovation and R&D Elasticities. Using the data we described above, we compute both the PDV-case and special case innovation elasticities for each firm (see Proposition 1). Figure 2 provides a comparison of the two sets of estimates. In particular, the figure shows a binned scatter plot of the special case innovation elasticities on the horizontal axis and the PDV innovation elasticities on the vertical axis. Each point in the figure represents one percent of firm-window observations in the sample. The dashed line visualizes the 45 degree line.

As can be seen, the vast majority of the distribution of innovation elasticities is very similar across the two approaches. In particular, with the exception of the top $2-3$ percent of firms with the highest innovation elasticities, both the special case and the PDV-based innovation elasticities are very similar.

Therefore, in what follows, we will use the special-case version of innovation elasticities as our preferred specification. Aside from the documented similarity to the PDV version, this is our preferred specification for two additional reasons. First, it requires fewer assumptions since it does not rely on estimating firms' discount rates and their respective present discounted values of R&D expenditures and profits. Second, it allows us to use the full sample of firms. In contrast, the Gormsen and Huber (2025) discount rates are available only for 49% of our sample.

The figure also shows that there is a large degree of heterogeneity in innovation elasticities across individual firms, further confirmed by Table 2. While the average innovation elasticity is about -0.6, the bottom decile of firms is characterized by an

Figure 1: Inverse of the R&D cost elasticity across industries



Note: The figure presents coefficients of regression (8) individually estimated in each 2-digit SIC sector (blue dots) and 95% confidence intervals (blue vertical lines). We present only those estimates that are statistically significantly different from zero. Observations are weighted with log patent citation count. Black dashed horizontal line corresponds to $\beta = 0.53$, i.e., the estimate in the first column in Table 1 with the shaded region indicating the corresponding 95% confidence interval.

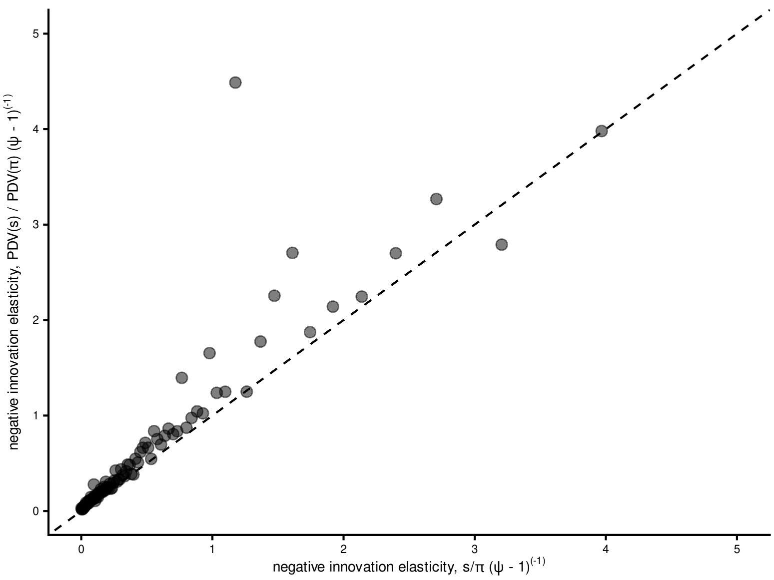
average innovation elasticity of -1.4. A similar degree of heterogeneity then carries over to firms' R&D elasticities, $\epsilon_i^s = \psi_i \epsilon_i$. These descriptive statistics already foreshadow some of the results of the next section – the ability to measure innovation elasticities at the firm-level has important aggregate implications.

2.3 Validation

In this sub-section, we validate our approach by comparing our sufficient statistics to estimates of R&D elasticities using established methods which utilize variation in R&D incentives across time, regions or eligibility thresholds. We do so by providing a novel empirical analysis of an R&D tax incentive change using Australian micro-data. At the end of this sub-section, we also show that our approach yields estimates that very closely line up with those in existing studies which use other datasets (Orbis and Compustat).

R&D Tax Incentives in Australia: Background and Data. Like many other countries, Australia provides R&D support to eligible businesses. Prior to 2012, businesses with sales exceeding AU$5 million were effectively eligible for an R&D subsidy of 37.5%, i.e. $P_s = 0.625$. In 2012, the government implemented a reform through which both the size threshold and the R&D subsidy changed. In particular, businesses with sales below

Figure 2: Innovation elasticities: PDV and special cases



The figure presents a binned scatter plot of the "special case" innovation elasticities – equation (7) – and the "PDV-based" innovation elasticities – equation (6). For visual clarity, we plot the negative of the elasticity and we use $\psi = 2$. Each dot represents 1 percent of the firm-window observations in our sample. The dashed line is the 45 degree line. To account for outliers, we restrict attention to firms with elasticities below 10 (in either of the two ways of computing elasticities). The median number of 5-year periods over which we calculate the PDV is four.

Table 2: Innovation and R&D elasticities: Descriptive Statistics

		Percentiles		
	Mean	10^{th}	50^{th}	90^{th}
Innovation elasticity, ϵ	-0.60	-1.44	-0.23	-0.03
R&D elasticity, $\epsilon^s = \psi\epsilon$	-1.02	-2.43	-0.40	-0.05

Notes: The table reports Compustat estimates of firms' innovation and R&D elasticities, ϵ and ϵ^s respectively, with sample selection described in the main text. The share of firm-window observations with reported R&D is 54%.

AU$20 million became eligible for a 45% subsidy ($P_s = 0.55$). By contrast, businesses with sales exceeding AU$20 million became eligible for a 40% subsidy ($P_s = 0.6$). More details on the institutional background can be found in Appendix D.3.

To analyze the effects of this policy change, we make use of Australia's Business Longitudinal Analysis Data Environment (BLADE). This dataset is based on administrative firm tax records provided by the Australian Bureau of Statistics and covers the universe of businesses registered for the Goods and Services Tax (value added tax) with total sales exceeding AU$75,000. It excludes sole traders or partnerships that submit personal income tax returns instead of business income tax returns. We use information on R&D expenditures, sales and profits for the years 2008-2016, around a key R&D policy change in 2012. In doing so, we focus on consolidated accounts – in line with the eligibility criteria of the R&D tax incentive.

In this policy setting, we estimate firms' average R&D elasticities, $\bar{\epsilon}^s$, using two different empirical specifications. First, we consider an event study design estimating firms' average R&D elasticities over time. Second, we directly test how the estimated average R&D elasticity relates to our sufficient statistics at the firm-level.[11]

R&D Tax Incentive in Australia: Event Study. Formally, we estimate the average R&D elasticity of treated firms (those with sales between AU$5 and AU$20 million) using the following event study design:

$$\log(R\&D)_{i,t} = \alpha + \sum_{t=2008,\neq 2011}^{2016} (\beta_t \mathbb{1}_{AU\$5-20M} \times \mathbb{1}_t + \delta_t \mathbb{1}_t) + \gamma \mathbb{1}_{AU\$5-20M} + \lambda X_{i,t} + u_{i,t}, \quad (9)$$

where $R\&D_{i,t}$ represents R&D expenditures of firm i in year t, $\mathbb{1}_{AU\$5-20M}$ is an indicator function equal to 1 for treated firms (those with sales between AU$5 million and AU$20 million in 2011 and 2012), $\mathbb{1}_t$ is an indicator function equal to 1 in year t, $X_{i,t}$ are other controls and $u_{i,t}$ are residuals. The coefficients of interest, β_t, measure the average percent difference in R&D expenditures between treated firms and the control group (those with sales exceeding AU$20 million) in year t relative to 2011.[12]

While estimates of β_t are interesting in their own right, our primary focus is on how these estimates compare to values implied by our methodology. Specifically, our theory predicts that – for a marginal change in the tax price of R&D – the coefficient β_{2012} would be equal to the theoretical average elasticity, $\bar{\epsilon}^s = -\frac{1}{N}\sum_i \frac{s_i}{\pi_i}\frac{\psi}{\psi-1}$, where the average is defined over the number of treated firms (N). Taking into account the actual change in the tax price of R&D, our theory predicts $\beta_{2012} = \bar{\epsilon}^s \Delta \log P_s$, where $\Delta \log P_s$ is the log change in subsidy rates of the treated group of firms relative to the control group.[13]

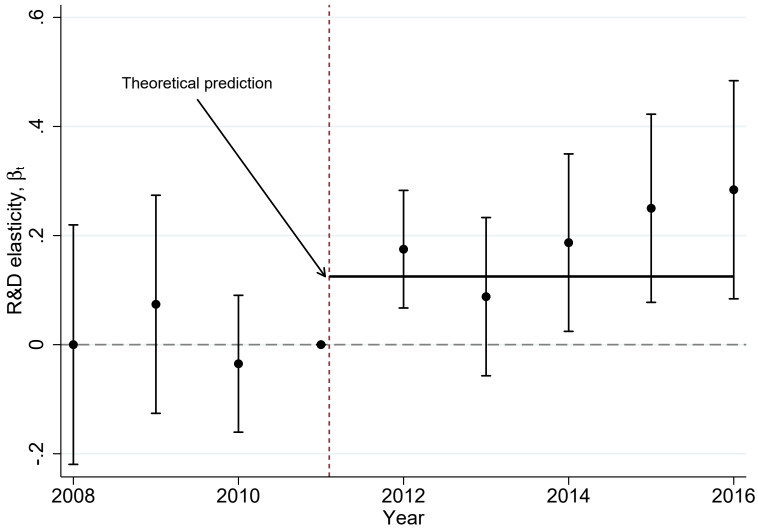
Figure 3 shows results of the event study estimation (9). As can be seen, while treated

[11]In both cases, we exclude firms with sales below AU$5 million, because these firms had an option – unobserved in the data – to exploit a second R&D incentive. To compute our sufficient statistics, we first need an estimate of firms' cost elasticities, ψ. Towards this end, we merge BLADE with the Intellectual Property Longitudinal Research Data (IPLORD), which provides annual information on firms' patents. Using this linked data, we estimate regression (8) for treated firms and obtain $\gamma = 0.396$ (standard error 0.135), implying an R&D cost elasticity of $\psi = 1/\gamma = 2.53$ – see Table A3. We use this value for the results that follow.

[12]Appendix D.7 demonstrates that our findings are robust to classifying the treatment and control group based on firms' pre-treatment sales in 2011, rather than excluding firms that change sales categories between 2011 and 2012. In addition, while the results from regression (9) represent estimates of the intention-to-treat (ITT) effect, Appendix D.6 shows similar results when employing a "fuzzy" DiD design to estimate the local average treatment effect (LATE) among "complying" firms.

[13]In the specific case of the policy change in 2012, this is equal to $\Delta \log P_s = \log(0.55/0.625) - \log(0.6/0.625) = -0.087$. In addition, in line with our theory, we use the pre-policy-change values (2011) of firms' R&D-to-profit ratio, s/π, to compute $\bar{\epsilon}^s$. Nevertheless, using time-varying averages changes little in our results.

Figure 3: R&D elasticity: Estimates and theoretical predictions



Note: The figure shows coefficient estimates of β_t from regression (9). For each estimate of β_t (the black dots), a 95% confidence interval is plotted (the black vertical lines). In addition, the horizontal line shows our "theoretical prediction", $\bar{\epsilon}^s = -\frac{1}{N}\sum_i \frac{s_i}{\pi_i}\frac{\psi}{\psi-1}\Delta \log P_s$.

and untreated firms had similar pre-trends, the policy change had a positive effect on the R&D expenditures of treated firms which persisted for several years. Importantly, our theoretical prediction lies well within the 95% confidence intervals.

R&D Tax Incentive in Australia: Direct Test of Sufficient Statistics. Next, we directly test whether the estimated R&D elasticities move one-for-one with our sufficient statistics at the firm-level. Towards this end, we estimate a familiar difference-in-differences design for the years 2011 and 2012 in which, however, we directly estimate heterogeneous treatment effects with respect to our sufficient statistics:

$$\begin{aligned}\log(R\&D)_{i,t} =& \alpha + (\beta_1 + \beta_2 e_i)\, \mathbb{1}_{AU\$5-20M} \times \mathbb{1}_{2012} \\ & + (\gamma_1 + \gamma_2 e_i)\, \mathbb{1}_{AU\$5-20M} + (\delta_1 + \delta_2 e_i)\, \mathbb{1}_{2012} + \eta e_i + \lambda X_{i,t} + u_{i,t}, \end{aligned} \quad (10)$$

where $e_i = \left(-\frac{s_i}{\pi_i}\frac{\psi}{\psi-1}\right) - \left(\frac{1}{N}\sum_i -\frac{s_i}{\pi_i}\frac{\psi}{\psi-1}\right)$ is the theoretical R&D elasticity in deviation from its mean.

The coefficient on the interaction term, β_2, directly measures the extent to which firms' R&D elasticities move (linearly) with our sufficient statistics. Our theory predicts that – for a marginal change in the tax price of R&D – the estimated R&D elasticity would move *one-for-one* with our theoretical elasticities, i.e. $\beta_2 = 1$. Taking into account the actual change in the tax price of R&D, our theory predicts $\beta_2 = 1 \times \Delta \log P_s$.

Table 3 reports the key results from our interacted DiD regression (10), together with the theoretical predictions. The table shows that we cannot reject the null hypothesis

Table 3: R&D elasticity and sufficient statistics

	(I)	(II)
Theoretical Prediction		
$\beta_2 = 1 \times \Delta \log P_s$	-0.087	-0.087
DiD Regression Results		
β_2, AU$5-20 mil. sales $\times$ 2012 $\times$ e	-0.106	-0.108
	(0.036)	(0.037)
Control for sales		yes
Observations	1784	1784

Notes: The table reports estimates of how firms' R&D elasticities co-move with our sufficient statistics, coefficient β_2 from regression (10), and the theoretical prediction. The regression sample includes firms in years 2011 and 2012 with sales exceeding AU$5 million, those that report positive R&D expenditure, positive operating profits and remain in the same sales category for both 2011 and 2012. Standard errors in parentheses are clustered at the firm level.

Table 4: Tax elasticity of R&D expenditures: Existing studies and our approach

	Appelt et al. (2025)				Hall (1993)
R&D elasticity, $\bar{\epsilon}^s$	All	Small	Medium	Large	Manuf.
Estimated	[-0.6,-0.4]	[-1.3,-0.9]	[-1.0,-0.8]	[-0.3,0.2]	[-1.5,-0.8]
Theoretical	-0.5	-1.2	-0.8	-0.3	-1.1

Notes: The table presents estimates of R&D elasticities. The top row reports the range of estimates of the elasticity of firms' R&D expenditures with respect to their price (tax) from Appelt et al. (2025) and Hall (1993), respectively. For Appelt et al. (2025), the values are based on Table 4 for all firms, and Tables 5 and 11 for size-based estimates (with and without dynamic adjustment). For Hall (1993), the values of the short-run elasticity are based on Tables 5a and 6, respectively. The bottom row uses the same Orbis country sample and Compustat sample, respectively, and the approach described in Proposition 1. "All" refers to all businesses in the sample, while "Small", "Medium" and "Large" are given by firms with 1-49, 50-249 and 250+ employees, following the definitions in Appelt et al. (2025). We assume that $\psi_i = 2$ for all i in Orbis and we use the estimated ψ for the manufacturing industry in Compustat. We only consider responsive firms, i.e., those with non-zero elasticities.

that $\beta_2 = 1 \times \Delta \log P_s$ (p-values of 0.622 and 0.594, respectively). Put together, results in this subsection suggest that not only do our theoretical predictions line up with the conventional event study estimates on average (Figure 3), they also correctly predict that in the data the estimated R&D elasticities move one-for-one with our sufficient statistics (Table 3).

Estimates of R&D Elasticities in Existing Studies. As a final step in the validation of our methodology, we also document that our sufficient statistics line up well against estimates in existing studies which use other datasets – Orbis and Compustat.

First, Appelt et al. (2025) estimate the average R&D elasticities, $\bar{\epsilon}^s$, using firm-level data and cross-country variation in R&D incentives in a sample of 11 OECD countries between the years 2000-2021.[14] Second, Hall (1993) obtains the average R&D elasticity from Compustat data on manufacturing firms between 1980 and 1991 by estimating an R&D investment Euler equation using instrumental variables. Employing the same datasets and sample periods, we compare estimates in the above two papers with our theoretical prediction, $\bar{\epsilon}^s = -\frac{1}{N}\sum_i \frac{s_i}{\pi_i}\frac{\psi}{\psi-1}$.

Table 4 shows the results. The first row shows estimates of R&D elasticities from the two studies and datasets. The second row shows our theoretical predictions, using the same datasets and sample periods. All the theoretical predictions fall within the estimated ranges, providing further validation for our approach.

These results show that R&D elasticities based on our sufficient statistics closely align with those estimated using established methodologies. The advantage of our approach is that we can measure R&D elasticities at the firm-level and without relying on variation in specific policies. We highlight these benefits next.

3 Aggregate Implications

We now turn to studying the aggregate implications of heterogeneity in firm-level innovation elasticities. This application showcases the type of analysis which is made possible by our methodology but which is not feasible with approaches that estimate "only" average innovation elasticities across large groups of firms.

In what follows, we first describe a baseline analytical framework and key *aggregate* objects of interest. The move from individual firms to their aggregation opens new questions such as the potential for spillovers between businesses, which we address through extensions to our baseline theory in Section 3.2. Then, we explain how the theory maps into the data and how we use the theoretical predictions to provide new facts about the aggregate implications of heterogeneity in firm-level R&D elasticities.

3.1 Theory: Baseline

We consider an environment in which heterogeneous firms i combine to produce aggregate nominal output, $Y = \sum_i y_i = \sum_i p_i q_i$, where p_i and q_i are the respective firm-level price

[14] Orbis data collects balance sheet information on both private and public companies across 200 countries and territories. The data – managed by Moody's – is sourced from business registers and private data providers and standardized across countries and over time. As before, we use consolidated company accounts to measure firms' R&D investment and profits. However, due to insufficient simultaneous coverage of firm-level R&D and patents across the selected sample of countries, we do not estimate country-specific R&D cost elasticities, ψ. Instead, we lean on the existing literature (see e.g. Hall and Ziedonis, 2001; Blundell et al., 2002) and use $\psi = 2$ for the Orbis validation exercise below. The country sample includes Australia, Belgium, Czechia, France, Italy, the Netherlands, New Zealand, Norway, Portugal, Slovakia and Sweden. The baseline estimates do not account for R&D incentive uptake since we do not have that information in our micro-data.

and quantities.

In this setting, real aggregate growth, G, is given by:

$$G = \sum_i m_i g_i, \tag{11}$$

where $m_i = y_i/Y$ is the market (sales) share of firm i and where $g_i = \mathrm{d}q_i/q_i$ is firm-level real output growth. Therefore, aggregate growth G is a weighted average of firm-level growth rates, where the weights are firms' sales shares (Domar weights).

At the same time, given the tax price of R&D, P_s, total government spending on R&D incentives is given by:

$$T = \sum_i \nu s_i, \tag{12}$$

where $\nu = (1 - P_s)/P_s$ measures the relative generosity of the R&D policy.

Bang for the Buck and its Components. We are now ready to define the aggregate objects of interest – the aggregate costs and benefits of changing the tax price of R&D.

Definition **1** (Bang for the Buck). *Consider a permanent change to the tax price of R&D,* $\mathrm{d}\log P_s$. *Then,*
(i) The Bang, B, is the associated impact response of the aggregate growth rate:

$$B = \frac{\mathrm{d}G}{\mathrm{d}\log P_s}. \tag{13}$$

(ii) The Buck, C, is the associated impact response of aggregate government spending on R&D support:

$$C = \frac{\mathrm{d}\log T}{\mathrm{d}\log P_s}. \tag{14}$$

(iii) The Bang for the Buck, A, is the ratio of the Bang and the Buck:

$$A = \frac{B}{C} = \frac{\mathrm{d}G/\mathrm{d}\log P_s}{\mathrm{d}\log T/\mathrm{d}\log P_s}. \tag{15}$$

The following proposition describes how the Bang and the Buck relate to firm-level variables. In anticipation of our empirical application, we also define the Bang and the Buck for mutually exclusive groups of firms Ω_k, where the set of all firms is given by $\Omega = \cup_k \Omega_k$.

Proposition **2** (The Bang and the Buck). *In the environment described above, the Bang* *and the Buck are given by:*

$$B = \sum_k B_k = \sum_{i\in\Omega} m_i g_i \epsilon_i, \tag{16}$$

$$C = \sum_k C_k = -\sum_{i\in\Omega} r_i \left(\frac{1}{\nu} - \epsilon_i^s\right), \tag{17}$$

where $r_i = \nu s_i / T$ *is the firm-level share in the government's expenditures on R&D incentives,* k *indicates mutually exclusive firm groups* Ω_k *and where* $B_k = \sum_{i\in\Omega_k} m_i g_i \epsilon_i$ *and* $C_k = -\sum_{i\in\Omega_k} r_i \left(\frac{1}{\nu} - \epsilon_i^s\right)$ *are, respectively, the group-specific Bang and Buck.*

The first part of Proposition 2 makes clear that the Bang is determined by a combination of firms' sales shares, m_i, their growth rates, g_i, and their innovation elasticities, ϵ_i. Intuitively, following a change in the tax price of R&D, individual firms can leave a mark on aggregate growth only if they are sufficiently large, fast-growing and responsive to the policy change.

The second part of Proposition 2 shows that the Buck is a weighted average of a direct $(1/\nu)$ and an indirect (ϵ_i^s) effect, where the weights reflect firms' shares in total government spending on R&D incentives, r_i. Intuitively, an R&D policy change will be particularly expensive if firms already receive a large chunk of R&D incentives, r_i, the policy is already relatively generous, $1/\nu$, and if firms' R&D responds strongly to the policy change, ϵ_i^s.

Put together, Proposition 2 highlights that the *aggregate* impact of changes in R&D incentives does not depend solely on firms' responsiveness (innovation or R&D elasticities). Instead, it is necessary to take into account the joint distribution of four firm-level elements: (i) R&D elasticities, ϵ_i^s, (ii) sizes, m_i, (iii) growth rates, g_i, and (iv) shares in government spending on R&D incentives, r_i. The following proposition makes explicit how heterogeneity in these firm-level factors drives the Bang and the Buck.

Proposition **3** (Components of the Bang and the Buck)**.** *(a) The Bang (16) can be expressed as:*

$$B = \sum_k B_k = \sum_k g_k m_k \epsilon_k \theta_k^g, \tag{18}$$

where each group-specific Bang, B_k*, is driven by four components:*

i) average (size-weighted) growth: $g_k = \sum_{i\in\Omega_k} \frac{y_i}{Y_k} g_i$

ii) size (sales share): $m_k = \sum_{i\in\Omega_k} y_i / Y = Y_k / Y$

iii) average innovation elasticity: $\epsilon_k = \frac{1}{N_k} \sum_{i\in\Omega_k} \epsilon_i$

iv) heterogeneity (in growth): $\theta_k^g = 1 + \frac{\text{cov}(g_k^y, \epsilon_k)}{g_k^y \epsilon_k}$,

where N_k *denotes the number of firms in group* k *and where* $g_k^y = \frac{1}{N_k} \sum_{i\in\Omega_k} y_i g_i$ *is the average output change in group* k*.*

(b) The Buck (17) can be expressed as:

$$C = \sum_k C_k = -\sum_k r_k \left(\frac{1}{\nu} - \epsilon_k^s \theta_k^s\right) \tag{19}$$

where each group-specific Buck, C_k, is driven by three components:

i) R&D "exposure": $r_k = \sum_{i\in\Omega_k} \nu s_i / T$

ii) average R&D elasticity: $\epsilon_k^s = \frac{1}{N_k}\sum_{i\in\Omega_k} \psi_i \epsilon_i$

iii) heterogeneity (in R&D): $\theta_k^s = 1 + \frac{\text{cov}(s_k, \epsilon_k^s)}{s_k \epsilon_k^s}$,

where $s_k = \frac{1}{N_k}\sum_{i\in\Omega_k} s_i$ is average R&D expenditure in group k.

Proposition 3 puts forward a convenient decomposition of the Bang and the Buck into average values of each of the components and the influence of firm heterogeneity. Intuitively, the Bang is higher if responsive firms (those with high $|\epsilon_i|$) are also large and/or fast growing businesses, i.e. $\theta^g > 1$. On the other hand, the Buck is higher if responsive firms are also R&D intensive, i.e. $\theta^s > 1$.

Note that equations (16) and (17) also naturally allow for firms which do not conduct any R&D. In this case, the innovation and R&D elasticities are simply equal to zero. In the following sub-section we consider the possibility of R&D spillovers across individual businesses.

Relation to Existing Concepts. Before moving on, we discuss how the Bang for the Buck relates to other concepts used in the literature. First, the Bang for the Buck is related to the "additionality ratio" which measures the dollars of additional R&D investment induced by a policy change, relative to the amount of forgone tax revenue.[15] In comparison, our Bang for the Buck (15) concept measures the impact on aggregate growth, rather than on private R&D spending. While closely related, the Bang for the Buck highlights that if policy makers are concerned with affecting aggregate growth, they need to account for a richer set of factors – not just firms' R&D responsiveness (R&D elasticity), but also their size and growth rates.

Second, our Bang relates to the "aggregate impact elasticity" in Atkeson and Burstein (2019). While the latter measures the impact on aggregate growth of a uniform increase in R&D expenditures across firms, the Bang relaxes precisely this and allows for heterogeneous R&D responses across firms to a given change in the tax price of R&D.

Finally, our "Buck" concept is closely related to the denominator of the marginal value of public funds (MVPF) of Hendren and Sprung-Keyser (2020). In the latter, the net cost

[15] The additionality ratio is sometimes also referred to as the bang for the buck or the cost effectiveness of R&D tax incentives (see e.g. Rao and Simcoe, 2026).

of a policy combines a mechanical outlay (our direct effect, $1/\nu$) with a behavioral "fiscal externality" (our indirect effect, ϵ^s). The Bang for the Buck $A = B/C$ therefore shares the MVPF's ratio structure – but focuses on aggregate growth, rather than private willingness to pay. In Section 4 we discuss how, with additional structure, our Bang would translate into consumption-equivalent welfare along the lines of Atkeson and Burstein (2019).

3.2 Theory: Extensions

In this subsection, we consider two extensions to our baseline framework: R&D spillovers between firms and firm-specific tax prices of R&D.

R&D Spillovers. To analyze how spillovers between firms may affect the Bang for the Buck, we consider the possibility that firm-level growth is a combination of growth driven by firms' "own" R&D efforts and growth driven by "external" spillovers from the R&D of other firms. Formally, we can then write:

$$g_i = \eta_{own} x_i + \eta_{ext} \underbrace{\sum_{j \neq i} \alpha_{i,j} x_j}_{S_i^{ext}}, \tag{20}$$

where η_{own} and η_{ext} represent the extent to which "own" and "external" innovation rates affect the growth rate of firm i and where $\alpha_{i,j}$ are "technology-proximity and quality-adjustment" weights. The latter summarize not only how closely firm i is related to firm j in terms of their technological fields, but also the quality of innovation in firm j. The following proposition describes the innovation elasticity and the Bang for the Buck under spillovers.

PROPOSITION 4 (Spillovers). *In the environment described above, innovation elasticities with spillovers can be expressed as*

$$\epsilon_i^{spill} = \omega_i \epsilon_i + (1 - \omega_i) \sum_{j \neq i} \sigma_{i,j} \epsilon_j, \tag{21}$$

where $\omega_i = \eta_{own} x_i / g_i$ is the fraction of "own-growth", ϵ_i is defined in (6) and $\sigma_{i,j} = \alpha_{i,j} x_j / S_i^{ext}$ are technology-proximity and quality-adjusted R&D shares of other firms, where $S_i^{ext} = \sum_{j \neq i} \alpha_{i,j} x_j$.
The Bang for the Buck is then given by

$$A^{spill} = \frac{\sum_i m_i g_i \epsilon_i^{spill}}{-\sum_i r_i \left(\frac{1}{\nu} - \epsilon_i^s \right)}. \tag{22}$$

First, Proposition 4 shows that the innovation elasticity of firm i which takes spillovers into account is a weighted average of the firm's individual innovation elasticity, ϵ_i, and that of all other businesses. However, innovation elasticities of all other businesses influence firm i only to the extent that they are technologically close, $\sigma_{i,j}$, and to the extent that external R&D drives some of firm i's growth, $1 - \omega_i$.

Next, Proposition 4 highlights that the spillover elasticity only enters the Bang, but not the Buck. The reason is that while R&D investment of a given firm may influence any other business, the government always subsidizes each firm only once.

Before moving on, let us consider the case when only firms of a particular group, Ω_k, are subsidized. The following proposition shows how we can decompose the group-specific Bang for the Buck into contributions of "own" R&D investment, internal and external spillovers and how all these relate to the case which ignores spillovers.

Proposition **5** (Bang for the Buck with & without Spillovers). *For firm group Ω_k, we can write the Bang for the Buck which accounts for spillovers as*

$$A_k^{spill} = A_k \left(\frac{B_k^{own}}{B_k} + \frac{B_k^{int} + B_k^{ext}}{B_k} \right), \tag{23}$$

where $A_k = B_k/C_k$ is the group-specific Bang for the Buck without spillovers with B_k and C_k defined in (18) and (19), respectively, and where $B_k^{own} = \sum_{i \in \Omega_k} m_i g_i \omega_i \epsilon_i$ is the "own" Bang, $B_k^{int} = \sum_{i \in \Omega_k} m_i g_i (1 - \omega_i) \sum_{j \in \Omega_k | \neq i} \sigma_{i,j} \epsilon_j$ is the "internal" Bang and $B_k^{ext} = \sum_{j \in \Omega_{\neq k}} m_j g_j (1 - \omega_j) \sum_{i \in \Omega_k} \sigma_{j,i} \epsilon_i$ is the "external" Bang.

The proposition makes clear how accounting for spillovers relates to the Bang for the Buck which ignores the effects firm innovation has on R&D of other businesses. Specifically, the "own" Bang is a scaled-down version of the Bang without spillovers, where the scaling depends on the extent to which firms' drive their own growth, ω_i. The internal Bang is a spillover effect from other firms within the subsidized group k. Finally, the external Bang quantifies how subsidizing firms in group k impacts growth of other firms outside the subsidized group due to technological spillovers.

Firm-Specific Tax Price of R&D. As an additional extension to our baseline framework, we consider that the aggregate effect of changes in R&D tax incentives is affected by possible differences in firm-specific levels of the tax price of R&D, $P_{i,s}$. This can occur in practice when R&D tax incentives differ across regions within a country or when they are a function of firms' corporate tax rates which themselves can vary with the level of firms' profits or sales.[16]

Concretely, we consider the impact of a common *proportional* change in firm-specific tax incentives in firm group k (which may include all firms), i.e. $d \log P_{i,s} = d \log P_s$

[16] An example of R&D incentives which depend on corporate tax rates is R&D "tax credits" which, for tax purposes, magnify a firm's actual R&D spending by a certain factor, effectively reducing its taxable profit (see e.g. Rao and Simcoe, 2026, for a discussion).

for all i. In this setting, firm-level innovation elasticities are unchanged relative to our baseline (6) and, therefore, the same is true for the Bang (16). In contrast, the following proposition shows how the Buck accounts for the heterogeneous generosity of R&D tax incentives.

PROPOSITION **6** (Firm-Specific Tax Price of R&D and the Buck)**.** *In the environment described above and with firm-specific levels of the tax price of R&D, $P_{i,s}$, the Buck is given by:*

$$C^{\tau} = \sum_k C_k^{\tau} = -\sum_{i \in \Omega} r_i \left(\frac{1}{\nu_i} - \epsilon_i^s \right), \tag{24}$$

where $\nu_i = (1 - P_{i,s})/P_{i,s}$.

The intuition behind Proposition 6 follows the same logic as our baseline results (see Proposition 2). In addition, however, the Buck explicitly reflects differences in firms' tax prices of R&D.

3.3 Measurement

In what follows, we exclusively use Compustat and the associated sample selection and data treatment described in Section 2.2. We now describe how the new concepts relating to aggregating individual firms and their innovation elasticities map into the data.

Sales and R&D Shares. Firm-level sales and R&D shares are given by $m_i = y_i/(\sum_i y_i)$ and $r_i = s_i/(\sum_i s_i)$, respectively, where we use Compustat `sales` and `xrd` to measure y_i and s_i.[17]

Growth Rates. Recalling that we use 5-year averaging windows, we use sales to measure growth as $g_i = y_{i,t}/y_{i,t-1} - 1$. For any firm-window observations with negative growth rates, we define innovation elasticities in the Bang as follows:

$$\epsilon_i^B = \begin{cases} \epsilon_i & \text{when } g_i \geq 0, \\ -\epsilon_i & \text{when } g_i < 0. \end{cases} \tag{25}$$

In this way, a decrease in the tax price of R&D boosts growth in businesses which are expanding and slows the contractions in firms with negative growth rates.[18]

Tax Price of R&D. For the baseline value of the tax price of R&D, P_s, we draw on OECD data which reports $P_s = 0.9$ for the U.S. economy (OECD, 2023). When

[17] Note that for a common tax price of R&D, $r_i = \nu s_i/T = \nu s_i/(\nu \sum_i s_i) = s_i/(\sum_i s_i)$.

[18] In Appendix D.12, we show that very similar results are obtained when dropping firms with negative growth rates altogether, in which case $\epsilon_i^B = \epsilon_i$. In addition, Appendix D.9 shows that the results are similar when considering employment growth instead.

considering the extension with firm-specific tax prices of R&D, we leverage state-specific R&D user cost data spanning 1985 to 2014 based on Wilson (2009); Falato and Sim (2014). In order to identify the geographical location of firms' R&D activity, we exploit the matched patent data. Specifically, for each patent we obtain the geographical location of all inventors of the patent. We aggregate all patents of a firm within a 5-year window and construct the geographical distribution of innovation activity of a firm.

To construct firm-level tax prices of R&D, we take the average of R&D tax credit values across all states, weighted by the share of innovators located in each state among all patents assigned to the firm within the time window. On average, we obtain tax prices of 0.88 with moderate heterogeneity over time and across firms: the inter-quartile range is three percentage points.

Spillovers. As discussed in Section 2.2, we use patent data collected by Kogan et al. (2017) which includes CPC technology classification codes. Following Bloom et al. (2013), we use CPC codes to measure closeness between firms in the technology space and to quantify R&D spillovers. Towards this end, we classify each patent using 3-digit CPC codes into one of 130 technology classes. Next, for each firm we compute the share of patents in all technology classes and compute the un-centered correlation between firm i's and firm j's patent shares, denoted by $\widetilde{\alpha}_{i,j}$.

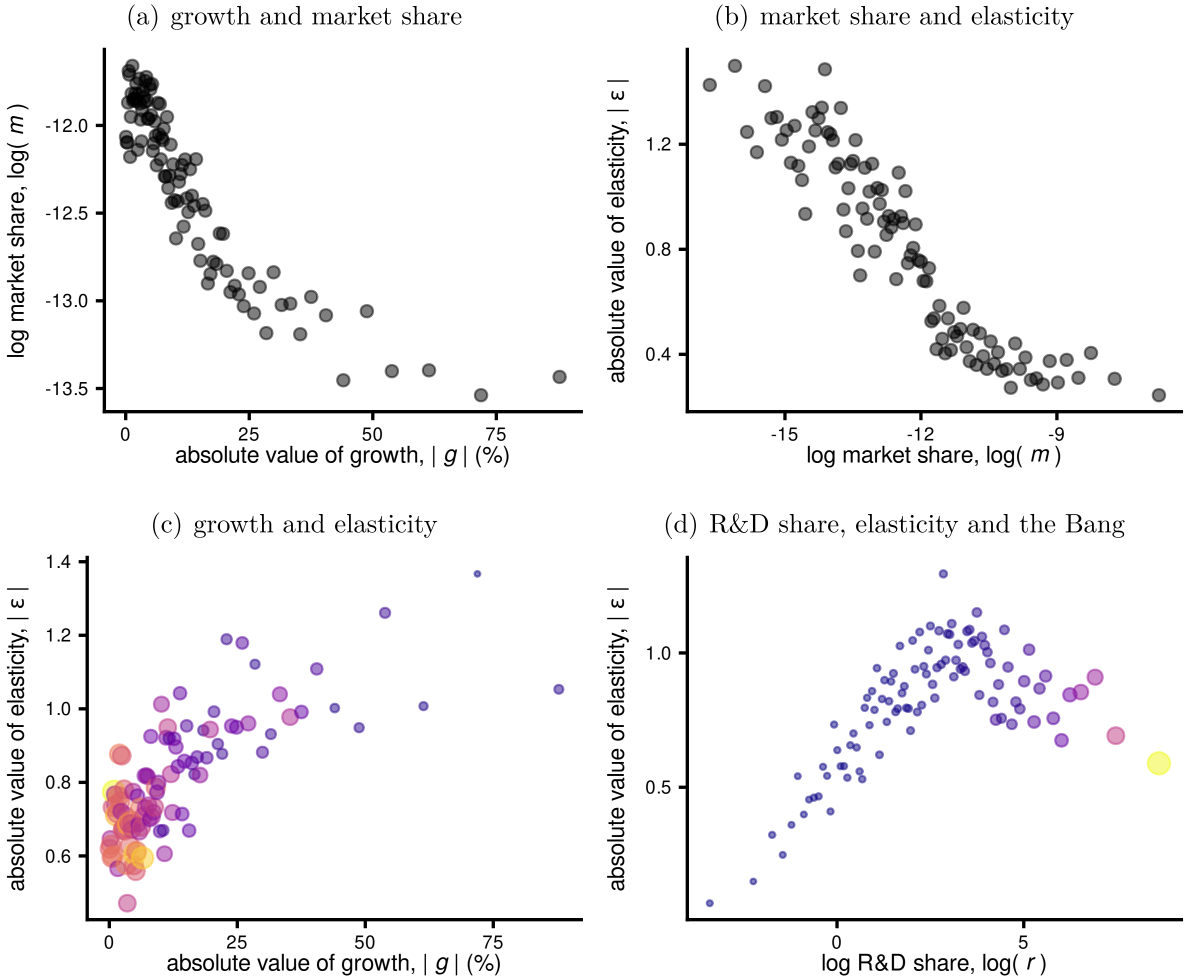
Finally, building on Griffith et al. (2011) and Bloom et al. (2013), we account for firms' importance in the patenting network by weighting each observation by the total patent citations accrued to the given firm. Specifically, let c_j denote all citations attributed to patents assigned to a given firm-period cell j. Then, our measure of the spillover benefiting firm i from firm j is given by the proximity-weighted, citation-adjusted R&D share $\sigma_{i,j} = \alpha_{i,j} \frac{x_j}{S_i^{ext}}$, where $\alpha_{i,j} = \widetilde{\alpha}_{i,j} \frac{c_j}{C}$ with $C = \sum_j c_j$ marking the total citation count and $S_i^{ext} = \sum_{j \neq i} \alpha_{i,j} x_j$. Intuitively, a firm i benefits from R&D expenses of a firm j if the two firms are patenting in the same set of CPC sectors and if firm j tends to generate patents with a high citation count.[19]

3.4 Results: Baseline

In this sub-section, we quantify the aggregate implications of *firm-level* heterogeneity in innovation elasticities, showcasing the key distinguishing factor of our approach relative

[19]Formally, $\widetilde{\alpha}_{i,j} = \frac{T_i T_j'}{\sqrt{T_i T_i'}\sqrt{T_j T_j'}}$, where T_i is a 130-element vector in which each element corresponds to the number of patents granted to firm i in a given CPC class. Further details on the patent data, its matching to Compustat and the construction of technology-proximity weights can be found in Appendix C.1. Note that firms' innovation rates are given by $x_j = s_j^{-1}(s_j(x_j))$, where the firm-specific function $s^{-1}(y) \propto y^{1/\psi}$ is the inverse of a firm's R&D cost function and is proportional to the firm's R&D expenditure to the power of $1/\psi$. Finally, in Appendix D.16, we document that adjusting spillovers for patent value estimated in Kogan et al. (2017), instead of citation counts, leads to very similar results.

Figure 4: Distribution of firm-level drivers of the Bang and the Buck



Note: The figure presents binned scatterplots for the components of the Bang. Each dot represents 1% of firm-window observations in Compustat data. In each panel, markers correspond to percentiles of the variable indicated on the horizontal axis. The corresponding value on the vertical axis is the unconditional sample average among firms within a given percentile. For the sake of visual clarity, we restrict attention to observations with growth rates less than 100% in absolute values and discard the top and bottom 2.5% of extreme observations in each variable.

to established methods.

Importance of Heterogeneity for the Bang and Buck. We begin by investigating the importance of firm heterogeneity for the *aggregate* Bang and Buck. As is described in Proposition 3, what matters for the magnitudes of the aggregate Bang and Buck is how innovation elasticities co-vary with firm size, R&D expenditure and growth rates. The impact of this cross-sectional variation is summarized by the heterogeneity terms θ^g and θ^s, respectively.

We begin by visualizing raw Compustat data to gauge the extent of heterogeneity in the key components of the Bang and the Buck. Figure 4 shows binned scatter plots for key components of the Bang for the Buck. In each panel of the figure, we consider percentiles of the distribution of the variable indicated on the horizontal axis and plot the corresponding sample average of the variables indicated on the vertical axis.

Panel (a) presents the distribution of (the absolute value of) growth rates, $|g|$, and firm sizes (market shares), m. The plot suggests a negative relationship between these variables. In particular, while larger firms are characterized by moderate growth rates, smaller businesses exhibit stronger growth. As we discuss in more detail below, this relationship is partly explained by firm age patterns, consistent with e.g. Haltiwanger et al. (2013).

Panel (b) shows a similar negative relationship between (the absolute value of) innovation elasticities, $|\epsilon|$, and firm size (market shares), m. Therefore, smaller firms tend to be relatively more responsive to R&D incentives.

Panel (c) then visualizes how the different components of the Bang interact. In particular, it plots the absolute values of innovation elasticities, $|\epsilon|$, against growth rates, $|g|$, with the size (and color) of each dot indicating the total market share, m. The larger the size (the warmer the color) of each dot, the larger is the market share of firms which find themselves in the given percentile of the growth rate distribution.

As can be seen, there is a positive relationship between growth rates and innovation elasticities. That said, as the size of the dots shows, most firms tend to be concentrated at the lower ends of both variables. The heterogeneity term, θ^g, summarizes all these joint patterns into a single contribution to the aggregate Bang, see Proposition 3. Quantitatively, in our baseline sample $\theta^g = 1.96$. Hence, our results suggest that existing approaches which are unable to measure innovation elasticities at the firm-level underestimate the impact of R&D subsidies on aggregate growth by one half.

Finally, Panel (d) presents the relationship between innovation elasticities, $|\epsilon|$, and R&D shares, r, while at the same time visualizing the magnitude of the Bang by the size (and color) of each percentile of firms. The larger (warmer colors of) the dots, the stronger the relative Bang of that percentile of firms compared to the average business. This panel, therefore, illustrates the key trade-off in the design of R&D subsidies: firms that are most responsive (high $|\epsilon|$) and are characterized by the largest Bang (large, warm-colored dots) are also the most expensive to support (high values of r). Quantitatively, this heterogeneity is summarized by $\theta^s = 0.98$, making the Buck which accounts for firm-level differences in R&D elasticities slightly lower compared to one computed with "just" an average elasticity across businesses. Therefore, if anything, these patterns further exacerbate the effect that accounting for firm heterogeneity has on the Bang for the Buck.

Relative Bangs for the Buck. We now move to analyzing the potential for R&D policies targeted at different groups of incumbent firms. We do so by focusing on the *relative* Bang for the Buck which measures how much more (or less) Bang for the Buck a policy maker could get if they targeted only firms in a particular group k, relative to simply handing out the subsidy indiscriminately to a random selection of businesses.

Towards this end, we define different groups of firms based on observable characteristics often considered in the literature: (i) R&D intensity, (ii) size, (iii) age and (iv) growth rates. In addition, we require that each group also accounts for exactly 50% of output ($m_k = 0.5$) so that all are comparable in their *capacity* to influence aggregate growth.[20]

Table 5 shows the results with the first three columns reporting the relative Bang for the Buck, the relative Bang and the relative Buck. The remainder of the table reports the components which drive the Bang and the Buck: (i) market shares, m_k, (ii) average innovation elasticities, $\bar{\epsilon}_k^B$ ($\bar{\epsilon}_k$), (iii) average sales-weighted growth rates, g_k, (iv) R&D shares, r_k, (v) heterogeneity in growth, θ_k^g and (vi) heterogeneity in R&D, θ_k^s.[21]

Relative Bangs for the Buck: R&D intensive firms. The first row of Table 5 shows that focusing innovation subsidies on R&D intensive firms provides a substantially *lower* Bang for the Buck than when simply handing out the subsidy indiscriminately to businesses, i.e. $A_{RD-int.}/A = 0.71$. Looking at the remaining columns reveals the reasons behind this.

While the relative Bang of R&D intensive firms is similar to that of a random selection of businesses, $B_{RD-int.}/B = 0.99$, the relative Buck is almost 40% higher. This is because R&D intensive firms account for almost 3/4 of all R&D expenditures. This out-sized "exposure" to innovation subsidies makes them an expensive policy target.

Relative Bangs for the Buck: Young firms. The second row of Table 5 shows that the Bang for the Buck of young businesses is almost 40% larger compared to that of a random selection of firms.

Looking at the underlying drivers reveals that the strong relative Bang for the Buck of young firms is predominantly driven by their higher responsiveness and stronger growth. These positive effects are somewhat dampened by the fact that young firms are relatively expensive as a policy target—their R&D share is slightly higher than their market share. Overall, however, young firms are more efficient in generating a Bang than they are costly in terms of their Buck.

Relative Bangs for the Buck: Small firms. Small firms seem to have very similar properties to young businesses, though slightly less favorable. At face value, this may be

[20] Appendix D.14 shows results when defining groups based on firm share splits, instead of sales shares.

[21] Recall from Section 3.3 that due to the presence of negative growth rates among some businesses, we separately report average innovation elasticities entering the Bang, $\bar{\epsilon}_k^B$ and those used in the Buck, $\bar{\epsilon}_k$. Note further that due to rounding of each Bang and Buck component, the products of the respective components do not exactly deliver the relative numbers reported in the first three columns. Finally, Appendix D.4 also considers asymptotic and bootstrapped standard errors for our relative Bangs for the Buck.

Table 5: Relative Bang for the Buck

Groups	Relative Bang & Buck A_k/A	B_k/B	C_k/C	Bang Components m_k	$\bar{\epsilon}_k^B$	g_k	θ_k^g	Buck Components r_k	$\bar{\epsilon}_k^s$	θ_k^s
R&D-int.	0.71	0.99	1.39	0.50	-0.31	0.04	2.49	0.73	-1.35	0.92
Young	1.38	1.52	1.10	0.50	-0.40	0.08	1.59	0.55	-1.45	1.14
Small	1.33	1.45	1.09	0.50	-0.37	0.06	2.22	0.55	-1.38	1.23
- Young	1.52	1.21	0.80	0.34	-0.40	0.08	1.93	0.39	-1.46	1.31
- Old	0.81	0.24	0.30	0.16	-0.04	0.02	30.78	0.16	-0.63	1.80
High-gr.	1.44	1.63	1.13	0.50	-0.81	0.13	0.52	0.58	-1.36	1.02
Random	1.00	1.00	1.00	0.50	-0.34	0.05	1.96	0.52	-1.35	0.98

Note: The first three columns of the table present Bangs for the Buck, Bangs and Bucks of different "groups" of firms, relative to those of a "random" selection of businesses (bottom row). Each group is defined to account for 50% of sales ($m_k = 0.5$), while conditioning on another observable. "R&D-intensive", "small", "young" and "high-growth" firms condition on, respectively, R&D-to-sales, sales, age and sales growth. The remaining columns report the components of the Bang and the Buck, see Proposition 3.

taken as evidence supportive of existing R&D schemes aimed at small and medium-sized enterprises.

However, size and age are closely related in the data (correlation coefficient of log sales and log age of 0.43). Moreover, as is well understood in the literature, size and age are not interchangeable (see e.g. Haltiwanger et al., 2013). In fact, as the fourth and fifth rows document, it is the prowess of young firms which props up the performance of small businesses.

In particular, about 2/3 of small firms are young ($m_{small-young} = 0.34$ vs $m_{small} = 0.5$) and these firms are characterized by strong growth and high innovation elasticities. In contrast, the 1/3 of small firms which are old display very low sensitivity to policies ($\bar{\epsilon}^B_{small-old} = -0.04$ vs $\bar{\epsilon}^B_{small} = -0.37$) and weak growth ($g_{small-old} = 0.02$ vs $g_{small} = 0.06$). This, despite a very favorable covariance between growth and responsiveness ($\theta^g_{small-old} = 30.8$ vs $\theta^g_{small} = 2.22$), drags the relative Bang for the Buck of small-old firms to just 81% of a random selection of businesses.

Therefore, according to our framework, age is the more direct indicator of a strong Bang for the Buck and would be a better dimension on which to base R&D policies. This finding parallels existing research on the importance of distinguishing firm age and size when analyzing job creation and productivity growth (see e.g. Haltiwanger et al., 2013). We contribute to this debate by pointing out that young businesses are also important propagators of R&D policies.

Relative Bangs for the Buck: High-Growth Firms. Finally, fast-growing businesses emerge as a similarly high-performing group of firms. The key reasons for their strong performance are high innovation elasticities ($\bar{\epsilon}^B_{high-gr} = -0.81$ vs $\bar{\epsilon}^B = -0.34$) and—by construction—strong growth rates ($g_{high-gr} = 0.13$ vs $g = 0.05$). This finding relates to work which highlights the importance of high-growth businesses for aggregate outcomes (see e.g. Haltiwanger et al., 2016).

These positive forces are dampened, but not outweighed, by the fact that high-growth firms are relatively expensive to subsidize ($m_{high-gr} < r_{high-gr}$) and that they are characterized by "unfavorable" heterogeneity. The latter can be seen from the fact that $\theta^g_{high-gr} < 1$, indicating that large/fast-growing firms typically have lower innovation elasticities.

3.5 Application: Extensions

As a final step in our analysis, we turn to investigating how extensions to our baseline specification may affect the conclusions discussed in the previous subsections. In particular, we consider the role of technological spillovers, firm-level differences in R&D incentives and dynamics.

Extensions: Firm-Specific R&D Incentives. Our main results are based on a common tax price of R&D. In this extension, we use the fact that in the U.S. the tax price of R&D differs across states and we estimate the effective tax price of R&D for individual firms, see Section 3.2.

The results are presented in the second column of Table 6. The first column repeats the relative Bang for the Buck from our baseline analysis and compares it to the case of firm-specific tax prices of R&D (second column). As can be seen, the results do not fundamentally change. In particular, the ranking of groups based on relative Bangs for the Buck remains unchanged.

Extensions: Spillovers. Next, we consider the impact of technological spillovers across businesses. Towards this end, we must first make a stand on how important spillovers are for firm-level growth on average. Existing research suggests that estimating technological spillovers across firms is an imprecise endeavor which is sensitive to model specifications, levels of aggregation or the particular empirical measure of spillovers (see e.g. Hall et al., 2010, for a survey). Therefore, in what follows we lean on recent estimates (see e.g. Matray, 2021; Dyevre, 2024) and offer a range of possible values for the importance of spillovers for firms' growth.

Columns 3-8 of Table 6 show the relative Bang for the Buck which accounts for spillovers. These values include a decomposition into the separate contributions of "own" R&D, "internal" spillovers within the considered group of firms and "external" spillovers which provide a boost for firms outside the group of businesses targeted by the policy – see Proposition 5.

Table 6: Relative Bang for the Buck: Firm-Specific Tax Prices of R&D and Spillovers

Firm group	$\frac{A_k}{A}$	$\frac{A_k^{P_{i,s}}}{A^{P_{i,s}}}$	Accounting for Spillovers					
			$\frac{A_k^{spill}}{A^{spill}}$	$\frac{B_k^{int}}{B_k}$	$\frac{B_k^{ext}}{B_k}$	$\frac{A_k^{spill}}{A^{spill}}$	$\frac{B_k^{int}}{B_k}$	$\frac{B_k^{ext}}{B_k}$
			$\omega = 0.85$			$\omega = 0.75$		
R&D-int.	0.71	0.67	0.91	0.39	0.35	1.01	0.64	0.58
Young	1.38	1.29	1.26	0.18	0.10	1.21	0.30	0.16
Small	1.33	1.23	1.21	0.14	0.13	1.15	0.24	0.21
High-gr.	1.44	1.39	1.37	0.25	0.07	1.34	0.42	0.12
Random	1.00	1.00	1.00	0.19	0.19	1.00	0.31	0.32

Note: The table presents the Bang for the Buck and its components when we account for firm-specific tax prices of R&D ($A_k^{P_{i,s}}/A^{P_{i,s}}$, 2nd column) and for technological spillovers between firms (columns 3-8). We use the sample of firms that we use for Table 5. A_k^{spill} and A_k correspond to Bang for the Buck with and without the spillovers, respectively. The components are defined in Proposition 5, where B_k marks the Bang without spillovers.

As can be seen from Table 6, R&D intensive businesses are the only firm group which features a stronger relative Bang for the Buck once spillovers are taken into account. This reflects our initial conjecture that R&D intensive firms may have positive growth effects stretching to other businesses (see columns B_k^{int}/B_k and B_k^{ext}/B_k). That said, the magnitude of these effects does not substantially alter our ranking of firm groups.

Extensions: Dynamics. Our baseline results provide the average effects across all firms and time periods. However, from the standpoint of the policy implementation, it is crucial to understand how the Bang for the Buck at the firm level changes over time. Are yesteryear's high-growth firms still the most cost-effective firms to support?

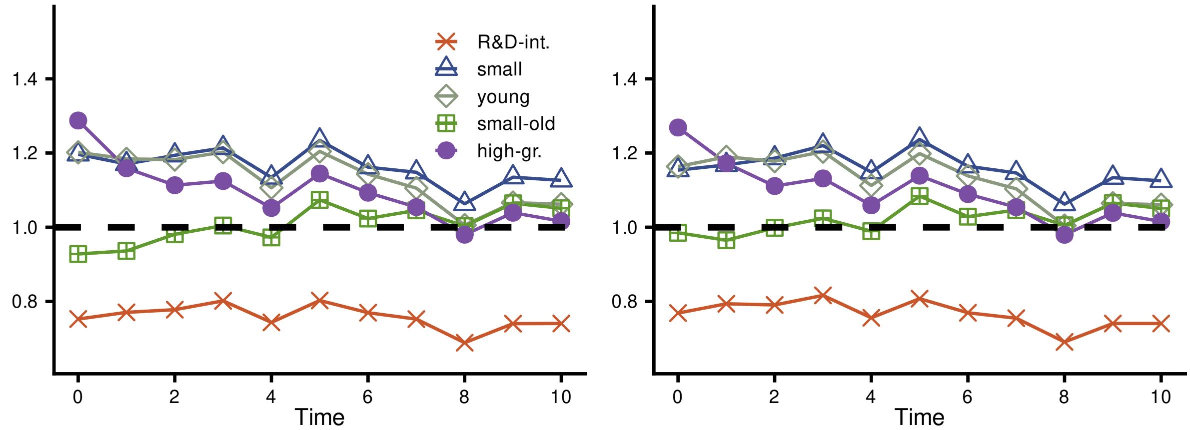
To account for possible changes in the relative Bangs for the Buck, we classify firms into our categories in the same fashion as described in Section 2.2 using annual data.[22] Then, keeping this classification fixed, we compute the *annual* relative Bang for the Buck for each group of firms over the following years and present the results in Figure 5. While the left panel (a) shows results for an unbalanced panel, the right panel (b) does the same for a balanced panel of firms. That is, we focus on firms that continuously report balance sheet information for at least 10 years.

Focusing on the unbalanced set of firms (left panel), the results show that relative Bangs for the Buck tend to converge (with the exception of those of R&D-intensive firms). However, young, and to a lesser extent fast-growing, firms retain their top ranking. As discussed previously, the performance of small firms is almost exclusively driven by the fact that many of them are young. Indeed, small-old firms have a relative Bang for the

[22] In the appendix we present the corresponding figure using averaged data over non-overlapping 5-year windows.

Figure 5: Relative Bangs for the Buck: Dynamics

(a) Relative Bangs for the Buck, Unbalanced Panel (b) Relative Bangs for the Buck, Balanced Panel



Note: The figure reports relative Bangs for the Buck across our firm groups using annual data. The classification into groups is made at $t = 0$. Then, we evaluate the changes in the Bangs and the Buck for a fixed classification. For example, the point on the purple line with circle markers at Time = 6 corresponds to the cost-effectiveness of supporting firms that were classified as high-growth 6 years ago. Left panel (a) shows results for pooled data of all firms. The right panel (b) restricts sample to a balanced panel of firms that continuously report balance sheet information for at least 10 years.

Buck which is persistently low. Similar patterns can be observed for the balanced panel, where conditioning on firm survival only exacerbates the prowess of young businesses.

The reason behind these patterns is that the Bang for the Buck does not depend on a single variable. Instead, it is determined by a combination of several drivers simultaneously (growth, size, R&D and innovation elasticities). Intuitively, as fast-growing firms expand quickly, they gain market share. Therefore, their weight in the economy remains relatively stable, even after their initially fast growth dissipates (see Appendix D.5 for further details).

4 Discussion

In this section, we briefly discuss the robustness of our findings along several dimensions and point towards possible fruitful avenues of future research.

4.1 Financial Frictions

Our analytical framework is based on firms optimally balancing the marginal costs and benefits of innovation. R&D subsidies directly affect the latter, since they make innovation cheaper. However, if firms are credit constrained, then R&D subsidies may have an additional impact on firms' decisions through alleviating credit constraints.

In this sense, our framework can be viewed as one aimed at the set of *unconstrained* firms. Arguably, this is the relevant group of businesses given the strong empirical skewness of R&D towards large firms. Indeed, Ottonello and Winberry (2024) develop a model of innovation under financial frictions and use Compustat data to estimate that indeed "the majority of innovation at a given time is performed by unconstrained firms."

Nevertheless, in Appendix B.4 we extend our framework to allow for financial constraints which are alleviated as firms grow larger. In this setting, firm-level innovation elasticities are a combination of two factors. First, and as in our baseline analysis, they depend on R&D-to-profits. Second, because of borrowing constraints, they also depend on the elasticity of the shadow value of funds with respect to R&D subsidies, ϵ_{i,χ,P_s}. It is worth distinguishing this second factor from the *level* effect of being constrained. A firm with a shadow value of funds above one innovates more than an otherwise identical unconstrained firm, precisely because growth relaxes its future constraints. The elasticity ϵ_{i,χ,P_s}, by contrast, captures the fact that a more generous subsidy itself relaxes the constraint, lowering the shadow value of funds and hence the marginal benefit of innovation. This second, marginal channel therefore *dampens* the R&D response of constrained firms.

While ϵ_{i,χ,P_s} is generally unobserved, it is also likely to be dwarfed by firms' innovation elasticities (R&D to profit ratios). To understand this, note that the shadow value of funds simply equals one for unconstrained firms. For constrained firms, Ottonello and Winberry (2020) estimate an average excess return on capital (a proxy for the shadow value of funds) of about 5%. Therefore, considering an extreme upper bound, where a marginal increase in R&D subsidies leads to complete *elimination* of all financing constraints, implies $\epsilon_{i,\chi,P_s} \approx 0.05$. This extreme upper bound is only about 8% of the *average* innovation elasticity in the data among firms reporting R&D expenses.

4.2 Technology Spillovers vs Business Stealing

In our approach, we focus on technological spillovers between businesses, measured by technology-proximity and quality-adjusted weights, $\sigma_{i,j}$. At the same time, however, firms may be connected also in the "product space." If so, businesses may encounter "business stealing" from firms which have managed to innovate, rather than reaping positive technological spillovers.

Indeed, according to estimates in Bloom et al. (2013), both types of spillovers (technology and product) are present in the data. However, technology spillovers are quantitatively much more important. Moreover, R&D intensive firms—which account for the majority of R&D expenditures—are characterized by the lowest relative Bang for the Buck (see Table 5). For these reasons, we focus on the possible positive effects of technological spillovers, as allowing for business stealing would likely only further decrease the contribution of R&D-intensive firms.

Similarly, there is evidence that R&D spillovers may be further amplified if they occur within a supply chain (see e.g. Li and Bosworth, 2020). Our framework could be extended to allow for additional spillover effects from firms connected via supply chain linkages, but it would require information on such business-to-business connections (see e.g. Bernard et al., 2022, for an analysis of firm-level data with production network information). We leave this avenue for future research.

4.3 Feedback Effects and the Distribution of Firms

Our baseline framework considers firms' partial equilibrium responses, holding aggregate growth fixed. While a full treatment of possible (general equilibrium) feedback effects would require imposing more structure on our environment, we now discuss how such feedback effects may impact our results.

First, survey evidence suggests that firm expectations respond much more strongly to firm-specific conditions, rather than aggregate developments (see e.g. Born et al., 2024). In this sense, our innovation elasticities are likely capturing the dominant trade-offs for incumbent firms.

Second, if firms ramp up their R&D efforts in response to more favorable tax incentives, aggregate innovation and growth will also increase. Stronger growth will, in turn, impact firm profits (values) through higher expected demand and factor prices and, ultimately, change firms' incentives to conduct R&D. However, to the extent that individual firms respond to such aggregate feedback effects in the same way, our conclusions about the *relative* Bangs for the Buck would remain unchanged.

Third, aside from their impact on incumbent firms—the topic of this paper—changes in R&D incentives may also affect the distribution of firms via firm entry and selection effects (see e.g. Acemoglu et al., 2018). While recent empirical evidence suggests that firm-level innovation does not seem to respond to R&D incentives along the extensive margin (see e.g. Dechezlepretre et al., 2023), a complete quantitative account of such effects would require a full general equilibrium model of endogenous growth. Our framework can, therefore, be viewed as an approximation around the prevailing firm distribution and associated balanced growth path, similar to e.g. Atkeson and Burstein (2019).

4.4 Real-Life Policies and Welfare Implications

Our methodology is based on marginal changes to firms' tax price of R&D. In reality, there are many specific details of policies, such as carry-forward/backward provisions or size-dependent corporate tax rates. While these characteristics are key for a correct measurement of the actual change in firms' tax price of R&D, $\Delta \log P_s$, they do not alter our novel sufficient statistics. In the validation exercise in Section 2.3, the economy-wide nature of the policy change, our use of 5-year averaging windows and the sample selection criteria all ensure that our theory aligns with the measurement in this regard.[23]

Finally, our analysis deliberately stops at aggregate growth as mapping growth to welfare would require imposing further model structure. It is nonetheless useful to make explicit how such a mapping would operate and what it would imply for the results presented in this paper.

In particular, converting the Bang (aggregate growth response) into a consumption-equivalent welfare gain would require taking into account the "social value of growth" and accounting for resource costs associated with a labor reallocation from production to innovation. However, as long as each firm inherits the same mapping to welfare, our results remain to hold. This is because we exclusively focus on firm-level results (Section 2) and *relative* aggregate results (Section 3).

Indeed, Atkeson and Burstein (2019) point to firm-level heterogeneity as a central open challenge. Our contribution is to measure precisely this – what are the heterogeneous growth responses of individual firms to R&D policies and how do they shape aggregate outcomes. In their terminology, we provide a mapping from individual firms to the aggregate growth impact elasticity. What we leave for future research is the possibility that in addition to firm-level heterogeneity in growth responses – the focus of this paper – there are differences across firms in the social value of growth or in the degree of intertemporal knowledge spillovers which Atkeson and Burstein (2019) stress are important for long-run welfare conclusions.

5 Conclusion

In this paper, we develop a sufficient-statistics approach to measuring R&D elasticities – how firms adjust R&D in response to tax incentives. Unlike existing methods, our approach identifies R&D elasticities for individual firms and without requiring policy variation across time, regions, or eligibility thresholds. We validate the resulting elasticities against estimates obtained from established difference-in-differences designs.

Applying the method to Compustat data reveals that firm heterogeneity is central to the aggregate effects of R&D policy. Accounting for differences in firms' R&D elasticities doubles the estimated aggregate growth effect of R&D incentives relative to a calculation based on the average elasticity. Young and fast-growing firms generate the largest increase in aggregate growth per dollar of subsidy, a result that remains robust when we allow for knowledge spillovers, firm-specific tax prices of R&D, and dynamic effects.

Our framework opens several avenues for future research. It can be used to examine whether different types of R&D generate different returns, to identify observable characteristics that predict firms' Bangs for the Buck, and to study how the ranking of firms

[23]The exercise does implicitly assume that our theoretical R&D elasticities – which are derived for marginal changes – carry over also to non-marginal policy reforms.

changes as the scale of an intervention increases. More broadly, it provides a tractable way to bring firm-level heterogeneity into the evaluation and design of innovation policy.

References

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William Kerr (2018) "Innovation, Reallocation and Growth," *American Economic Review*, 108 (11), 3450–3491.

Acemoglu, Daron and Dan Cao (2015) "Innovation by Entrants and Incumbents," *Journal of Economic Theory*, 157, 255–294.

Aghion, Philippe and Peter Howitt (1992) "A Model of Growth through Creative Destruction," *Econometrica*, 60, 323–351.

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

Akcigit, Ufuk, John R. Grigsby, Tom Nicholas, and Stefanie Stancheva (2022a) "Taxation and Innovation in the 20th Century," *Quarterly Journal of Economics*, 137 (1), 329–385.

Akcigit, Ufuk, Douglas Hanley, and Stefanie Stancheva (2022b) "Optimal Taxation and R&D Policies," *Econometrica*, 90 (2), 645–684.

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

Appelt, Silvia, Matěj Bajgar, Chiara Criscuolo, and Fernando Galindo-Rueda (2025) "How effective are R&D tax incentives? Reconciling the micro and macro evidence," CEP Discussion Paper, 2071.

Atkeson, Andrew and Ariel Burstein (2019) "Aggregate Implications of Innovation Policy," *Journal of Political Economy*, 127 (6), 2625–2683, 10.1086/701823.

Bernard, Andrew, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes (2022) "The Origins of Firm Heterogeneity: A Production Network Approach," *Journal of Political Economy*, 130, 1765–1804.

Bloom, Nicholas, Rachel Griffith, and John Van Reenen (2002) "Do R&D Tax Credits Work?" *Journal of Public Economics*, 85 (1), 1–31.

Bloom, Nicholas, Mark Schankerman, and John van Reenen (2013) "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 81 (4), 1347–1393.

Bloom, Nicholas, John van Reenen, and Heidi Williams (2019) "A Toolkit of Policies to Promote Innovation," *Journal of Economic Perspectives*, 33 (3), 163–184.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer (2002) "Individual Effects and Dynamics in Count Data Models," *Journal of Econometrics*, 108 (1), 113–131.

Born, Benjamin, Zeno Enders, Manuel Menkhoff, Gernot J. Müller, and Knut Niemann (2024) "Firm Expectations and News: Micro v Macro," CEPR Discussion Paper 17768.

Cavenaile, Laurent, Murat Alp Celik, Pau Roldan-Blanco, and Xu Tian (2021) "Style over Substance? Advertising, Innovation, and Endogenous Market Structure," mimeo.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suarez Serrato, and Daniel Xu (2021) "Notching R&D Investment with Corporate Income Tax Cuts in China," *American Economic Review*, 111 (7), 2065–2100.

Chetty, Raj (2009) "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451–488.

Dechezlepretre, Antoine, Elias Einio, Ralf Martin, Kieu-Trang Nguyen, and John van Reenen (2023) "Do Tax Incentives Increase Firm Innovation? An RD Design for R&D, Patents and Spillovers," *American Economic Journal: Economic Policy*.

Dyevre, Arnaud (2024) "Public R&D Spillovers and Productivity Growth," mimeo.

Falato, Antonio and Jae Sim (2014) "Why Do Innovative Firms Hold So Much Cash? Evidence from Changes in State R&D Tax Credits," *FEDS Working Paper*.

Gormsen, Niels and Kilian Huber (2025) "Corporate Discount Rates," *American Economic Review*, 115 (6), 2001–2049.

Griffith, Rachel, Sokbae Lee, and John Van Reenen (2011) "Is distance dying at last? Falling home bias in fixed-effects models of patent citations," *Quantitative Economics*, 2 (2), 211–249, https://doi.org/10.3982/QE59.

Guceri, Irem and Li Liu (2019) "Effectiveness of Fiscal Incentives for R&D: Quasi-Experimental Evidence," *American Economic Journal: Economic Policy*, 11 (1), 266–291.

Hall, Bronwyn (1993) "R&D tax policy during the 1980s: success or failure," *Tax Policy and the Economy*, 1–35.

Hall, Bronwyn, Jacques Mairesse, and Pierre Mohnen (2010) "Measuring the Returns to R&D," Handbook of the Economics of Innovation, Chapter 24.

Hall, Bronwyn and John van Reenen (2000) "How effective are fiscal incentives fo R&D? A review of the evidence," *Research Policy*, 29, 449–469.

Hall, Bronwyn and Rosemarie Ziedonis (2001) "The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979-1995," *RAND Journal of Economics*, 101–128.

Haltiwanger, John, Ron Jarmin, and Javier Miranda (2013) "Who Creates Jobs? Small Versus Large Versus Young," *Review of Economics and Statistics*, 95 (2), 347–361.

Haltiwanger, John, Ron S. Jarmin, Robert Kulick, and Javier Miranda (2016) "High Growth Young Firms: Contribution to Job, Output, and Productivity Growth," in *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, 11–62: University of Chicago Press.

Hendren, Nathaniel and Ben Sprung-Keyser (2020) "A Unified Welfare Analysis of Government Policies," *Quarterly Journal of Economics*, 135 (3), 1209–1318.

Hopenhayn, Hugo A. and Richard Rogerson (1993) "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101 (5), 915–938, 10.1086/261909.

Ignaszak, Marek and Petr Sedláček (forthcoming) "Customer Acquisition, Business Dynamism and Aggregate Growth," *Review of Economic Studies*.

Klette, Tor Jakob and Samuel Kortum (2004) "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112 (5), 986–1018.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman (2017) "Technological Innovation, Resource Allocation, and Growth," *The Quarterly Journal of Economics*, 132 (2), 665–712, 10.1093/qje/qjw040.

Lentz, Rasmus and Dale T Mortensen (2008) "An empirical model of growth through product innovation," *Econometrica*, 76 (6), 1317–1373, Publisher: Wiley Online Library.

Li, Yuxin and Derek Bosworth (2020) "R&D Spillovers in a Supply Chain and Productivity Performance in British Firms," *The Journal of Technology Transfer*, 45, 177–204.

Luttmer, Erzo GJ (2007) "Selection, growth, and the size distribution of firms," *The Quarterly Journal of Economics*, 122 (3), 1103–1144.

——— (2010) "Models of growth and firm heterogeneity," *Annu. Rev. Econ.*, 2 (1), 547–576.

Matray, Adrien (2021) "The Local Innovation Spillovers of Listed Firms," *Journal of Financial Economics*, 141 (2), 395–412.

Mukoyama, Toshihiko and Sophie Osotimehin (2019) "Barriers to Reallocation and Economic Growth: The Effects of Firing Costs," *American Economic Journal: Macroeconomics*, 11 (4), 235–270.

OECD (2023) "OECD R&D tax incentives database," data retrieved on 30.08.2023, `https://www.oecd.org/innovation/tax-incentives-RD-innovation/`.

Ottonello, Pablo and Thomas Winberry (2020) "Financial Heterogeneity and the Investment Channel of Monetary Policy," *Econometrica*, 88 (6), 2473–2502.

——— (2024) "Capital, Ideas, and the Costs of Financial Frictions," mimeo.

Rao, Nirupama (2016) "Do Tax Credits Stimulate R&D Spending? The Effect of R&D Tax Credit in its First Decade," *Journal of Public Economics*, 140, 1–12.

Rao, Nirupama and Timothy Simcoe (2026) "Tax Incentives for Research and Development: Policy Design and Evidence," in Jones, Benjamin and Josh Lerner eds. *Entrepreneurship and Innovation Policy and the Economy*, 5: University of Chicago Press.

de Souza, Gustavo (2025) "R&D Subsidy and Import Substitution: Growing in the Shadow of Protection," Working Paper Series WP 2023-37, Federal Reserve Bank of Chicago.

Sterk, Vincent ⓡ Petr Sedláček ⓡ Benjamin Pugsley (2021) "The Nature of Firm Growth," *American Economic Review*, 111 (2).

Wilson, Daniel J (2009) "Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of R&D tax credits," *The Review of Economics and Statistics*, 91 (2), 431–436.

Online Appendix

Firm Heterogeneity and the Aggregate Effects of R&D Policy

Marek Ignaszak, Daniel Robbins and Petr Sedláček

Contents

A Proofs

This Appendix provides proofs for all the propositions in the main text.

A.1 Proposition 1

We start from frameworks in which profits (firm value) are proportional to output, $v_{i,t} = q_{i,t}\widehat{v}_{i,t}$ and $\widehat{v}_{i,t}$ is independent of $q_{i,t}$ (see e.g. Klette and Kortum, 2004; Lentz and Mortensen, 2008; Acemoglu and Cao, 2015; Mukoyama and Osotimehin, 2019). In this case, we can write firm value as:

$$\widehat{v}_{i,t} = \max_{x_{i,t}} \widehat{\pi}_{i,t} + \beta_i \frac{q_{i,t+1}}{q_{i,t}} \widehat{v}_{i,t+1} = \max_{x_{i,t}} \widehat{\pi}_{i,t} + \beta_i(1+g_{i,t})\widehat{v}_{i,t+1}. \tag{A1}$$

The optimal R&D decision is then given by

$$\psi_i \frac{\widehat{s}_{i,t}}{x_{i,t}} = \beta_i \lambda_i \widehat{v}_{i,t+1} \tag{A2}$$

$$\psi_i \frac{s_{i,t}}{x_{i,t}} = \beta_i \frac{\lambda_i}{1+g_{i,t}} v_{i,t+1}, \tag{A3}$$

where we use the fact that profits explicitly include R&D expenditures, $\pi_{i,t} = y_{i,t} - w_{i,t} - s_{i,t}$, with $y_{i,t}$ being sales and $w_{i,t}$ being all other operational costs which are, by construction, independent of R&D. While the above implicitly defines optimal innovation rates, we are interested in understanding how such R&D decisions change in response to a change in the tax price of R&D, P_s. Towards this end, we totally differentiate (A2) with respect to dP_s:[1]

$$\overbrace{\psi_i \frac{\widehat{s}_{i,t}}{x_{i,t}}}^{MC} \epsilon_i(\psi_i - 1)\frac{dP_s}{P_s} = \tag{A4}$$

$$-\overbrace{\beta_i\lambda_i\widehat{v}_{i,t+1}}^{MB} \frac{\sum_{j=1} \widetilde{\beta}_{i,t+j}^{j-1} \left(\overbrace{\frac{\widehat{s}_{i,t+j}}{P_s}}^{\text{direct effects}} - \epsilon_i \overbrace{\frac{\psi_i \widehat{s}_{i,t+j} - \beta_i\lambda_i x_{i,t+j}\widehat{v}_{i,t+1+j}}{P_s}}^{\text{indirect effects} = 0} \right)}{\widehat{v}_{i,t+1}} dP_s$$

$$\epsilon_{i,t} = -\frac{\sum_{j=1}^{\infty} \beta_i^{j-1} s_{i,t+j}}{v_{i,t+1}} \frac{1}{\psi_i - 1}, \tag{A5}$$

where $\widetilde{\beta}_{i,t} = \beta_i(1+g_{i,t})$ and $\widehat{v}_{i,t+1} = \sum_{j=1} \widetilde{\beta}_{i,t+j}^{j-1} \widehat{\pi}_{i,t+j}$ and where indirect effects via future R&D expenditures cancel because of the envelope theorem. Notice also that in the special

[1] We evaluate the resulting elasticity at a baseline level $P_s = 1$. In addition, we ignore the direct effect on marginal costs in the current period. This is consistent with the elasticity representing firms' response to a policy change which takes place from the next period onwards.

case of $s_{i,t}/\pi_{i,t} = s_i/\pi_i$ for all t, the innovation elasticity becomes

$$\epsilon_i = -\frac{s_i}{\pi_i}\frac{1}{\psi_i - 1}. \tag{A6}$$

A.2 Proposition 2

For convenience, let us repeat the key definitions:

$$\begin{aligned}
B =& \frac{\mathrm{d}G}{\mathrm{d}\log P_s}, && \text{(A7)}\\
C =& \frac{\mathrm{d}\log T}{\mathrm{d}\log P_s}, && \text{(A8)}\\
T =& (1-P_s)\sum_i \frac{s_i(x_i)}{P_s}, && \text{(A9)}\\
G =& \sum_i m_i g_i, && \text{(A10)}
\end{aligned}$$

where we note that both the Bang, B, and the Buck, C, are defined as *impact* responses, holding firms' sales shares fixed. In this setting, the Bang is given by differentiating (A10) with respect to P_s:

$$\begin{aligned}
dG =& \sum_i m_i \frac{\partial g_i}{\partial P_s} dP_s = \sum_i m_i \epsilon_i \frac{g_i}{P_s} dP_s = \sum_i m_i g_i \epsilon_i \frac{dP_s}{P_s}\\
B =& \frac{dG}{d\log P_s} = \sum_i m_i g_i \epsilon_i,
\end{aligned}$$

where we have used the definition of our innovation elasticity, $\epsilon_i = \frac{\partial x_i}{\partial P_s}\frac{P_s}{x_i} = \frac{\partial g_i}{\partial P_s}\frac{P_s}{g_i}$.

Next, to obtain the Buck, we differentiate (A9) with respect to P_s, using that $dP_s = -d(1-P_s)$:

$$\begin{aligned}
dT =& \sum_i \left(\overbrace{-\frac{s_i(x_i)}{P_s}}^{\text{direct effect}} + \overbrace{(1-P_s)\frac{\partial \frac{s_i(x_i)}{P_s}}{\partial x_i}\frac{\partial x_i}{\partial P_s}}^{\text{indirect effect}} \right) dP_s = -\sum_i \left(\frac{s_i(x_i)}{P_s} - (1-P_s)\psi_i \epsilon_i \frac{\frac{s_i(x_i)}{P_s}}{P_s} \right) dP_s\\
=& -\sum_i (1-P_s)\frac{s_i(x_i)}{P_s}\left(\frac{P_s}{1-P_s} - \psi_i\epsilon_i\right)\frac{dP_s}{P_s} = -T\sum_i \overbrace{\frac{(1-P_s)\frac{s_i(x_i)}{P_s}}{T}}^{=r_i}\left(\overbrace{\frac{P_s}{1-P_s}}^{=1/\nu} - \overbrace{\psi_i\epsilon_i}^{=\epsilon_i^s} \right)\frac{dP_s}{P_s}\\
C =& \frac{d\log T}{d\log P_s} = -\sum_i r_i\left(\frac{1}{\nu} - \epsilon_i^s\right).
\end{aligned}$$

A.3 Proposition 3

Recall that the Bang and the Buck are given by:

$$B = \sum_i m_i g_i \epsilon_i,$$
$$C = -\sum_i r_i \left(\frac{1}{\nu} - \epsilon_i^s\right),$$

where $m_i = y_i/Y$ and $r_i = \nu s_i/T$. We can now rewrite the Bang as

$$B = \frac{1}{Y}\sum_i y_i g_i \epsilon_i = \frac{N}{Y}\frac{1}{N}\sum_i g_i^y \epsilon_i = \frac{1}{\overline{y}}\left(\overline{g}^y \overline{\epsilon} + \mathrm{cov}(g^y, \epsilon)\right)$$

where $g_i^y = y_i g_i$ and where we have used the fact that for two variables x and y, $\frac{1}{N}\sum_{i=1}^N x_i y_i = \overline{x}\,\overline{y} + \mathrm{cov}(x, y)$. Next, denoting $n_k = N_k/N$ and $\overline{y} = Y/N$, we use the same logic for group-specific Bangs:

$$\begin{aligned}
B_k =& \frac{1}{Y}\sum_{i\in\Omega_k} y_i g_i \epsilon_i = \frac{1}{Y}\frac{N_k}{N_k}\sum_{i\in\Omega_k} g_i^y \epsilon_i = \frac{N_k}{Y}\overline{g}_k^y \overline{\epsilon}_k \left(1 + \frac{\mathrm{cov}(g_k^y, \epsilon_k)}{\overline{g}_k^y \overline{\epsilon}_k}\right) \\
=& \frac{N_k}{Y}\frac{1}{N_k}\sum_{i\in\Omega_k} y_i g_i \overline{\epsilon}_k \left(1 + \frac{\mathrm{cov}(g_k^y, \epsilon_k)}{\overline{g}_k^y \overline{\epsilon}_k}\right) = \frac{Y_k}{Y}\sum_{i\in\Omega_k}\frac{y_i}{Y_k} g_i \overline{\epsilon}_k \left(1 + \frac{\mathrm{cov}(g_k^y, \epsilon_k)}{\overline{g}_k^y \overline{\epsilon}_k}\right) \\
=& m_k g_k \overline{\epsilon}_k \theta_k^g.
\end{aligned}$$

From the above, we see that $B = \sum_k g_k m_k \overline{\epsilon}_k \theta_k^g$. Using the same logic, and noting that with a common tax incentive $r_i = \nu s_i/T = s_i/S$, we can rewrite the Buck as

$$C = -\frac{1}{\nu} + \frac{1}{S}\sum_i s_i \epsilon_i^s = -\frac{1}{\nu} + \frac{N}{S}\frac{1}{N}\sum_i s_i \epsilon_i^s = -\frac{1}{\nu} + \left(\overline{\epsilon}^s + \frac{\mathrm{cov}(s, \epsilon^s)}{\overline{s}}\right),$$

where $\overline{\epsilon}^s = \frac{1}{N}\sum_i \psi_i \epsilon_i$ is the average R&D elasticity and $\overline{s} = S/N$ is average R&D expenditure. Extending the above to group-specific Bucks, we can write

$$\begin{aligned}
C_k =& -\frac{1}{S}\sum_{i\in\Omega_k} s_i \left(\frac{1}{\nu} - \epsilon_i^s\right) = -\frac{1}{\nu} r_k + \frac{1}{S}\frac{N_k}{N_k}\sum_{i\in\Omega_k} s_i \epsilon_i^s = -\frac{1}{\nu} r_k + \frac{N_k}{S}\overline{s}_k \overline{\epsilon}_k^s \left(1 + \frac{\mathrm{cov}(s_k, \epsilon_k^s)}{\overline{s}_k \overline{\epsilon}_k^s}\right) \\
=& -\frac{1}{\nu} r_k + \frac{N_k}{S}\frac{1}{N_k}\sum_{i\in\Omega_k} s_i \overline{\epsilon}_k^s \left(1 + \frac{\mathrm{cov}(s_k, \epsilon_k^s)}{\overline{s}_k \overline{\epsilon}_k^s}\right) = -\frac{1}{\nu} r_k + r_k \overline{\epsilon}_k^s \left(1 + \frac{\mathrm{cov}(s_k, \epsilon_k^s)}{\overline{s}_k^s \overline{\epsilon}_k^s}\right) \\
=& -r_k \left(\frac{1}{\nu} - \overline{\epsilon}_k^s \theta_k^s\right).
\end{aligned}$$

From the above, we see that $C = -\sum_k r_k \left(\frac{1}{\nu} - \overline{\epsilon}_k^s \theta_k^s\right)$.

A.4 Proposition 4

Let us begin by repeating how a firm's growth depends on its "own" R&D efforts and on "external" spillovers:

$$g_i = \eta_{own} x_i + \eta_{ext} \underbrace{\sum_{j \neq i} \alpha_{i,j} x_j}_{S_i^{ext}} .$$

In this environment, changing subsidies affects firms' own incentives to conduct R&D, but it also creates spillover effects from increased R&D of other firms:

$$\begin{aligned}
\epsilon_i^{spill} = \frac{\partial g_i}{\partial P_s} \frac{P_s}{g_i} &= \left[\eta_{own} \frac{\partial x_i}{\partial P_s} + \eta_{ext} \sum_{j \neq i} \alpha_{i,j} \frac{\partial x_j}{\partial P_s} \right] \frac{P_s}{g_i} \\
&= \left[\eta_{own} \frac{x_i}{P_s} \epsilon_i + \eta_{ext} \sum_{j \neq i} \alpha_{i,j} \epsilon_j \frac{x_j}{P_s} \right] \frac{P_s}{g_i} \\
&= \underbrace{\eta_{own} \frac{x_i}{g_i}}_{\omega_i} \epsilon_i + \underbrace{\eta_{ext} \frac{S_i^{ext}}{g_i}}_{1-\omega_i} \sum_{j \neq i} \underbrace{\frac{\alpha_{i,j} x_j}{S_i^{ext}}}_{\sigma_{i,j}} \epsilon_j = \omega_i \epsilon_i + (1 - \omega_i) \sum_{j \neq i} \sigma_{i,j} \epsilon_j ,
\end{aligned}$$

Substituting ϵ_i^{spill} into the definition of the Bang, we can write

$$B^{spill} = \sum_i m_i g_i \epsilon_i^{spill} .$$

Note that the Buck is unchanged because it is defined as the change in government expenditures, which are not affected by spillovers.

A.5 Proposition 5

Let us consider that only group k of firms is subsidized and that spillover effects exist. In this case, we can write the group-specific Bang as:

$$B_k^{spill} = \overbrace{\sum_{i \in \Omega_k} m_i g_i \omega_i \epsilon_i}^{\text{"own" Bang, } B_k^{own}} + \overbrace{\sum_{i \in \Omega_k} m_i g_i (1 - \omega_i) \sum_{j \in \Omega_k | \neq i} \sigma_{i,j} \epsilon_j}^{\text{"internal" spillover Bang, } B_k^{int}} + \overbrace{\sum_{j \in \Omega_{\neq k}} m_j g_j (1 - \omega_j) \sum_{i \in \Omega_k} \sigma_{j,i} \epsilon_i}^{\text{"external" spillover Bang, } B_k^{ext}}$$

Note that the group-specific Buck is the same as without spillovers. Next, we can write the group-specific Bang for the Buck with spillovers as follows:

$$\begin{aligned}
A_k^{spill} = &\frac{B_k^{spill}}{C_k} = \frac{B_k}{C_k} \left(\frac{B_k^{own}}{B_k} + \frac{B_k^{int}}{B_k} + \frac{B_k^{ext}}{B_k} \right) \\
= &A_k \left(\frac{B_k^{own}}{B_k} + \frac{B_k^{int} + B_k^{ext}}{B_k} \right) \qquad \text{(A11)}
\end{aligned}$$

A.6 Proposition 6

With heterogeneous tax prices of R&D, total government expenditures on R&D incentives are given by $T = \sum_i (1 - P_{i,s}) \frac{s_i(x_i)}{P_{i,s}}$. We now consider a common percent change in R&D tax incentives, i.e. $d\log P_{i,s} = d\log P_s$ for all i. In this case, individual innovation rates and the Bang remain the same as in our baseline. However, the Buck now accounts for heterogeneous levels of firms' tax prices of R&D. In particular, the Buck is now given by

$$\begin{aligned}
dT =& \sum_i \left(\overbrace{-\frac{s_i(x_i)}{P_{i,s}}}^{\text{direct effect}} + \overbrace{(1-P_{i,s}) \frac{\partial \frac{s_i(x_i)}{P_{i,s}}}{\partial x_i} \frac{\partial x_i}{\partial P_{i,s}}}^{\text{indirect effect}} \right) dP_{i,s} = -\sum_i \left(\frac{s_i(x_i)}{P_{i,s}} - (1-P_{i,s}) \psi_i \epsilon_i \frac{\frac{s_i(x_i)}{P_{i,s}}}{P_{i,s}} \right) dP_{i,s} \\
=& -\sum_i (1-P_{i,s}) \frac{s_i(x_i)}{P_{i,s}} \left(\frac{P_{i,s}}{1-P_{i,s}} - \psi_i \epsilon_i \right) \frac{dP_{i,s}}{P_{i,s}} \\
=& -T \sum_i \overbrace{\frac{(1-P_{i,s}) \frac{s_i(x_i)}{P_{i,s}}}{T}}^{=r_i} \left(\overbrace{\frac{P_{i,s}}{1-P_{i,s}}}^{=1/\nu_i} - \overbrace{\psi_i \epsilon_i}^{=\epsilon_i^s} \right) \frac{dP_{i,s}}{P_{i,s}} \\
C^\tau =& \frac{d\log T}{d\log P_s} = -\sum_i r_i \left(\frac{1}{\nu_i} - \epsilon_i^s \right),
\end{aligned}$$

where the last line uses the fact that $d\log P_{i,s} = d\log P_s$.

B Additional Analytical Results

In this Appendix, we describe an example of an entire structural model consistent with our framework in the main text as well as provide additional extensions to our baseline framework.

B.1 Workhorse Model

We now present an example of a full structural model that is consistent with our framework in the main text. Time is discrete and we use primes to denote next period values.

Production. We assume that there is a continuum of individual firms, indexed by i, each producing a differentiated final consumption good, q_i. The final goods are consumed by the representative household endowed with constant elasticity preferences over a consumption bundle,

$$U = \sum_{t=0}^{\infty} \beta^t u(C_t), \qquad C = \left[\int_i q_i^{\frac{\eta-1}{\eta}} \, \mathrm{d}i \right]^{\frac{\eta}{\eta-1}},$$

where $\eta > 1$ is the elasticity of substitution between goods. The household faces an aggregated budget constraint

$$Z' + \int_i c_i p_i \, \mathrm{d}i = WN + (1+R)Z,$$

where Z marks the holdings of a diversified equity portfolio of all firms in the economy, p_i is the firm-specific goods price (relative to the aggregate price index P which is normalized to 1), W is the aggregate wage and N is the aggregate labor supply. Without loss of generality, we assume that the household supplies inelastically one unit of labor.

The optimal consumption choice implies that each firm faces a downward-sloping demand for its product:

$$q_i = p_i^{-\eta} Y, \tag{A12}$$

where Y is aggregate expenditure. Final goods are produced using a linear technology combining production labor, n_i, and a firm-specific (endogenous) productivity, a_i:

$$q_i = a_i n_i. \tag{A13}$$

Innovation. Firms strive to improve their own productivity by investing resources into R&D. Firms invest s_i units of labor into R&D in return for a probability x_i with which they successfully improve upon their current productivity level. We assume that R&D costs are convex in the innovation probability, x:

$$s_i = \overline{s}_i x_i^{\psi_i}, \tag{A14}$$

where we assume that $\psi_i > 1$ and $\overline{s}_i > 0$. If successful, innovations lead to an increase in firm-level productivity by a factor of $(1 + \lambda_i)$, where $\lambda_i > 0$ is a potentially firm-specific constant:

$$a_i' = \begin{cases} a_i(1+\lambda_i) & \text{with probability } x_i \\ a_i & \text{with probability } 1 - x_i. \end{cases} \tag{A15}$$

Optimality conditions. Let us now describe the optimality conditions characterizing firms' choice of the price, production employees and R&D employees, $p(a_i)$, $n(a_i)$ and $s(a_i)$, respectively, as a function of firms' productivity a_i. The current profit is given by

$$\pi(a_i) = p(a_i)^{1-\eta} Y - Wn(a_i) - Ws(a_i).$$

Let β_i be a firm-specific discount factor that also includes an exogenous exit probability. The outside option of exiting firms is normalized to zero. Exiters are replaced by entrants who draw initial productivity from the distribution of incumbent firms. The firm value function becomes

$$V_i(a_i) = \pi(a_i) + \beta_i \left[x_i V_i(a_i(1+\lambda_i)) + (1-x_i)V_i(a_i)\right]$$

where we highlight that firm value functions are allowed to be firm-specific due to the heterogeneity in the deep parameters governing the technology and discount factors.

The optimal pricing choice is standard and implies a constant markup over the marginal cost, $p(a_i) = \frac{\eta}{\eta-1}\frac{W}{a_i}$. Consequently, the production labor demand is $n(a_i) = p_i^{-\eta} Y/a_i = \left(\frac{\eta}{\eta-1}\right)^{-\eta} a_i^{\eta-1} W^{-\eta} Y$. Finally, the optimal R&D labor satisfies

$$\psi_i W \overline{s}_i x_i^{\psi_i - 1} = \beta_i \left(V(a_i(1+\lambda_i)) - V(a_i)\right)$$

Equilibrium. Let $Q = \left[\int a_i^{\eta-1}\,\mathrm{d}i\right]^{\frac{1}{\eta-1}}$ be the aggregate productivity index. We restrict attention to balanced growth path equilibria in which all aggregate variables grow at the same rate as Q. In what follows, let us denote stationarized variables by hats, for example $\widehat{a} = a/Q$ marks stationarized productivity.

Now, we introduce a key assumption that specifies the shape of the R&D cost function $\overline{s}_i \equiv \rho_i \left(\widehat{a}_i\right)^{\eta-1}$, where $\rho_i > 0$ is a potentially firm specific, but time-invariant constant. Below, we show that this is the natural scaling consistent with the existence of a balanced growth path equilibrium (BGP). Before that, let us denote the stationarized profits as

$$\widehat{\pi}(\widehat{a}_i) = \frac{\pi(a_i)}{Q} = \widehat{a}_i^{\eta-1} \widehat{W}^{1-\eta} \widehat{Y} \left(\frac{\eta}{\eta-1}\right)^{-\eta} \frac{1}{\eta-1} - \widehat{W} \rho_i \widehat{a}_i^{\eta-1} x_i^{\psi_i}$$

To obtain analytical solutions, we restrict $\psi_i = 2$ for all i. Then, the optimality conditions specified above imply that firm value is proportional to $\widehat{a}_i^{\eta-1}$. We prove this claim using a guess and verify method. Guess that the value function satisfies $V \equiv Q\widehat{v}_i \widehat{a}_i^{\eta-1}$ for a potentially firm-specific but time invariant constant $\widehat{v}_i > 0$. Given this guess, the optimal R&D policy implies that

$$x_i^* = \frac{\beta_i}{2\widehat{W}\rho_i} \frac{(1+g)^{\frac{1}{\eta-1}}}{1+g} \widehat{v}_i \left((1+\lambda_i)^{\eta-1} - 1\right)$$

The resulting x_i^* is firm-specific, but—crucially—time invariant as it does not depend on the productivity $\widehat{a}_i$. With these intermediate results, the definition of firm value implies

$$\widehat{v}_i \widehat{a}_i^{\eta-1} = \widehat{a}_i^{\eta-1} \left[\widehat{W}^{1-\eta} \widehat{Y} \mu^* - \widehat{W} \rho_i (x_i^*)^2 + \beta_i \frac{(1+g)^{\frac{1}{\eta-1}}}{1+g} \widehat{v}_i \left(x_i^*(1+\lambda_i)^{\eta-1} + 1 - x_i^*\right) \right] \tag{A16}$$

where $\mu^* = \left(\frac{\eta}{\eta-1}\right)^{-\eta} \frac{1}{\eta-1}$ and where we used our guess on the functional form of the value function which we subsequently verified.

Along a BGP equilibrium, all aggregate variables grow at the rate $1+g$ equal to the growth rate of aggregate productivity index $Q = \left[\int a_i^{\eta-1}\,\mathrm{d}i\right]^{\frac{1}{\eta-1}}$. To verify this observe that since all costs are denominated in labor units and aggregate price index is normalized to unity, the aggregate resource constraint reads $C = Y$. Note further that the aggregate consumption can be written as $C = \left[\int_i \left(\left(\frac{\eta}{\eta-1}\right)^{-\eta} a_i^{\eta} W^{-\eta} C\right)^{\frac{\eta-1}{\eta}} \mathrm{d}i\right]^{\frac{\eta}{\eta-1}}$, which implies that $W = \frac{\eta-1}{\eta}\left[\int a_i^{\eta-1}\,\mathrm{d}i\right]^{\frac{1}{\eta-1}} = \frac{\eta-1}{\eta}Q$. Consequently, $\widehat{W} = \frac{\eta-1}{\eta}$.

Next, the labor market clearing implies that $N = 1 = \int n(a_i)\,\mathrm{d}i + \int s(a_i)\,\mathrm{d}i$. Using the optimal policy derived above, we can show that $\int n(a_i)\,\mathrm{d}i = CQ^{-1}$. Furthermore, $\int s(a_i)\,\mathrm{d}i = \int \rho_i \left(\frac{a_i}{Q}\right)^{\eta-1}(x_i^*)^2\,\mathrm{d}i$ is stationary. Therefore, the labor demand is constant along the balanced growth path.

Let us note that the labor market clearing requires that the labor devoted to R&D does not grow in a balanced growth equilibrium. This is only feasible if the mass of researchers required to deliver a given innovation probability grows at the same rate as the market size. Otherwise, as firms' profits increase, firms would spend increasing amounts of resources on R&D which violates labor market clearing in the limit. The postulated scaling of the R&D cost is therefore the natural way to satisfy this requirement.

Model-Implied Innovation Elasticities. Finally, we show that the stylized model above satisfies our key results. First, observe that expected firm level growth rate is

$$1+g = \frac{(a_i')^{\eta-1}}{a_i^{\eta-1}} = 1 + x_i^*[(1+\lambda_i)^{\eta-1} - 1].$$

Therefore, $\log g = \log x_i^* + \log[(1+\lambda_i)^{\eta-1} - 1]$. With these intermediate results we have

$$\log s_i = \log \rho_i + \log\left(\widehat{a}_i\right)^{\eta-1} + \psi(\log g - \log[(1+\lambda_i)^{\eta-1} - 1])$$

Consequently, $\frac{\mathrm{d}\log s_i}{\mathrm{d}\log g} = \psi_i$. It then follows that

$$\frac{s_i}{\pi(a_i)} = \frac{s_i/Q}{\pi(a_i)/Q} = \frac{\widehat{W}\rho_i x_i^2}{\widehat{W}^{1-\eta}\widehat{Y}\left(\frac{\eta}{\eta-1}\right)^{-\eta}\frac{1}{\eta-1} - \widehat{W}\rho_i x_i^2}.$$

The above expression is independent of the firm's idiosyncratic state a_i. Furthermore, all variables on the right-hand side are constant along the balanced growth path. Consequently, while firm specific, the ratio of R&D expenses to profits is time invariant. Therefore, the model described in this section adheres to the special case of Proposition 1, where $\epsilon_i = -s_i/\pi_i$ since $\psi_i = 2$.

B.2 Innovation Elasticities in Klette and Kortum (2004)

Equation (3) in Klette and Kortum (2004) defines firm value as:

$$rV(n) = \max_{I} \{\widetilde{\pi} n - C(I, n) + I\,[V(n+1) - V(n)] - \mu n\,[V(n) - V(n-1)]\}$$

where n is the number of product lines, r is the interest rate, $I(n) = \lambda n$ is the optimal innovation rate with λ being the innovation intensity, $C(I(n), n) = nc(\lambda)$ are the associated costs of R&D, π are profits per product line and μ is the endogenous creative destruction rate. As can be seen, this structure implies $V(n) = vn$ (i.e. firm value is proportional to the number of product lines) which in turn gives

$$rv = \widetilde{\pi} - c(\lambda) + \lambda v - \mu v$$

Collecting terms, we have

$$0 = \widetilde{\pi} - c(\lambda) + (\lambda - r - \mu)v.$$

In the notation of "our" framework, we have operational profits, π_o, given by $\widetilde{\pi}$, R&D costs s_i given by $c(\lambda)$ and the innovation rate x_i given by λ, with a continuation value of $v(\lambda - r - \mu)$ where the latter is effectively $\beta/(1+G)(1+g)$ in our setting.

Let us now assume a particular functional form $c(\lambda) = P_s c_o \lambda^{\psi}$, where $P_s = 1$ in the Klette and Kortum setting. The optimality condition for innovation investment is given by:

$$c'(\lambda) = \psi \frac{c(\lambda)}{\lambda} = v = \frac{\widetilde{\pi} - c(\lambda)}{r + \mu - \lambda}.$$

We now take the total differential of the above w.r.t. P_s (ignoring the direct effect on marginal costs as before), where $\epsilon = \frac{\partial \lambda}{\partial P_s}\frac{P_s}{\lambda}$ is the innovation elasticity defined in (4). First, the marginal cost side:

$$\begin{aligned}
&\frac{\psi}{\lambda}\left(\psi \frac{c(\lambda)}{\lambda} \overbrace{\frac{\partial \lambda}{\partial P_s}\frac{P_s}{\lambda}}^{=\epsilon} \frac{\lambda}{P_s} - \frac{c(\lambda)}{\lambda} \overbrace{\frac{\partial \lambda}{\partial P_s}\frac{P_s}{\lambda}}^{=\epsilon} \frac{\lambda}{P_s}\right) dP_s \\
=& \frac{\psi c(\lambda)}{\lambda}\left(\psi\epsilon - \epsilon\right)\frac{dP_s}{P_s} = \underbrace{\psi \frac{c(\lambda)}{\lambda}}_{=MC} \epsilon(\psi - 1)\frac{dP_s}{P_s}
\end{aligned}$$

Second, the marginal benefit part:

$$-\frac{v}{\widetilde{\pi}-c(\lambda)}\left[\overbrace{\frac{c(\lambda)}{P_s}}^{\text{direct eff.}}+\psi\frac{c(\lambda)}{\lambda}\epsilon\frac{\lambda}{P_s}\right]dP_s+\frac{v}{r+\mu-\lambda}\epsilon\frac{\lambda}{P_s}dP_s$$

$$=-\underbrace{v}_{=MB}\frac{c(\lambda)}{\widetilde{\pi}-c(\lambda)}\frac{dP_s}{P_s}-\epsilon v\underbrace{\left[\psi\frac{c(\lambda)}{\widetilde{\pi}-c(\lambda)}-\frac{\lambda}{r+\mu-\lambda}\right]}_{=0\text{ (envelope theorem)}}\frac{dP_s}{P_s}$$

Putting the above two together and canceling marginal costs and benefits on each side, we get

$$\epsilon=-\frac{1}{\psi-1}\underbrace{\frac{c(\lambda)}{\widetilde{\pi}-c(\lambda)}}_{=s/\pi}.$$

The above shows that in the canonical Klette and Kortum (2004) model, firms' innovation elasticities are characterized precisely by our sufficient statistic, ϵ_i.

B.3 Stochastic Profits

In this section, we discuss the extension in which the firm profits follow a stochastic process, rather than a deterministic path as in the baseline framework in the main text. As before, we assume that firm profits follow an additive specification in operating profits $\widetilde{\pi}_{i,t}=\widetilde{\pi}^o_{i,t}-s_{i,t}$. Here $\widetilde{\pi}^o_{i,t}$ denotes a realization of an i.i.d. random variable with a finite first moment. Any such random variable can always be represented by $\widetilde{\pi}^o_{i,t}=\pi^o_{i,t}+\zeta_{i,t}$ where $\pi^o_{i,t}$ is the mean of the process and $\zeta_{i,t}$ is a zero-mean, i.i.d. stochastic process. Therefore, we can represent the stochastic process for profits $\widetilde{\pi}$ as the sum of a deterministic component, π, and a stochastic component, ζ.

Consider a firm that maximizes *expected* present discounted value of all future profits:

$$v_i(q_{i,t})=\max_{x_{i,t}}\mathbb{E}_t\sum_{j=0}\beta_i^j\widetilde{\pi}_i(q_{i,t+j})$$

where $\mathbb{E}_t$ marks the expectation conditional on period-t information set. In this setting, all our results carry over as firms decide based on expected values, the stochastic component of profits is additive and $\mathbb{E}_t\sum_{j=0}\beta_i^j\zeta_{t+j}=0$.

Consequently, the optimal R&D investment (firm-level growth) satisfies the following optimality condition:

$$\psi_i\frac{s_{i,t}}{x_{i,t}}=\frac{\partial v_{i,t}}{\partial x_{i,t}},\tag{A17}$$

which is the same condition as (3) – noting that firm value is in expectation. Following the same steps as in Section A.1 delivers the following expression for innovation elasticities

$$\epsilon_i = -\frac{\sum_{j=1}^{\infty} \mathbb{E}_t \beta_i^{j-1} s_{i,t+j}}{v_{i,t+1}} \frac{1}{\psi_i - 1}$$

B.4 Financial Frictions

In this Appendix, we extend our baseline framework to allow for financial (or other) frictions, e.g. along the lines of Ottonello and Winberry (2024). We do so by considering a "wedge", $\chi_{i,t} \geq 1$, which enters the condition for optimal innovation decisions. In this setting, equation (3) becomes:

$$\psi_i \frac{s_{i,t}}{x_{i,t}} = \chi_{i,t} \beta_i \frac{\lambda_i}{1 + g_{i,t}} v_{i,t+1}, \tag{A18}$$

In the terminology of Ottonello and Winberry (2024), $\chi_{i,t}$ represents the shadow value of funds. For unconstrained firms, $\chi_{i,t} = 1$ and our baseline framework applies. For constrained firms, $\chi_{i,t} > 1$ and firms benefit from innovation not only because growth increases future profits, but also because it loosens financing constraints. In particular, $\frac{\partial \chi_{i,t}}{\partial q_{i,t}} \leq 0$. For the same reason, cheaper R&D brought about by more generous subsidies also leads to a loosening of financing constraints, $\epsilon_{i,\chi,P_s} = \frac{\partial \chi_{i,t}}{\partial P_s} \frac{P_s}{\chi_{i,t}} \geq 0$.

Next, following the same steps as in Appendix A.1 and assuming that $\epsilon_{i,\chi,P_s} \geq 0$ is constant over time though possibly heterogeneous across firms, we can derive firms' innovation elasticities as:[2]

$$\overbrace{\psi_i \frac{\widehat{s}_{i,t}}{x_{i,t}}}^{MC} \epsilon_i (\psi_i - 1) \frac{dP_s}{P_s} =$$

$$\underbrace{\chi_{i,t} \beta_i \lambda_i \widehat{v}_{i,t+1}}_{MB} \left(\overbrace{\frac{\epsilon_{i,\chi,P_s}}{P_s}}^{\text{effect via the wedge}} - \frac{\sum_{j=1}^{\infty} \widetilde{\beta}_{i,t+j}^{j-1} \left(\overbrace{\frac{\widehat{s}_{i,t+j}}{P_s}}^{\text{direct effects}} - \epsilon_i \overbrace{\frac{\psi_i \widehat{s}_{i,t+j} - \chi_{i,t+j} \beta_i \lambda_i x_{i,t+j} \widehat{v}_{i,t+1+j}}{P_s}}^{\text{indirect effects} = 0} \right)}{\widehat{v}_{i,t+1}} \right) dP_s$$

$$\epsilon_{i,t} = -\frac{1}{\psi_i - 1} \left(\frac{\sum_{j=1}^{\infty} \beta_i^{j-1} s_{i,t+j}}{v_{i,t+1}} - \epsilon_{i,\chi,P_s} \right),$$

[2]Throughout, we assume that firms take the shadow value of funds, $\chi_{i,t+j}$, as given when choosing $x_{i,t}$, i.e. they do not internalize how their own R&D affects their future shadow value. Without this assumption, $\partial \chi_{i,t+j} / \partial x_{i,t} \neq 0$ – since $\partial \chi_{i,t} / \partial q_{i,t} \leq 0$ – and the indirect effects below would not cancel through the envelope theorem.

Once again, the special case when $s_{i,t}/\pi_{i,t}$ (and ϵ_{i,χ,P_s}) is constant at the firm-level boils down to

$$\epsilon_i = -\frac{1}{\psi_i - 1}\left(\frac{s_i}{\pi_i} - \epsilon_{i,\chi,P_s}\right). \tag{A19}$$

C Further Details on Data

In this Appendix, we provide further details on the nature of the three datasets we use in our analysis: Compustat, BLADE and Orbis. In addition, we provide more details on our empirical results and several robustness exercises.

C.1 Compustat and Patent Data

As discussed in detail in Kogan et al. (2017), the inventor of a patented innovation can assign the granted property rights to another legal entity, for example a corporation. Therefore, granted patents may have, in addition to the inventor, an assignee, that is, one or more corporations or persons. Kogan et al. (2017) match corporate assignees of all U.S. patents to publicly traded U.S. corporations whose stock market returns can be found in the CRSP database. Finally, using the CRSP identifiers, we match the corresponding balance sheet information in the Compustat dataset. The resulting information allows us to calculate the number of patents for each firm in Compustat.

Following Bloom et al. (2013); Jaffe (1986), we use Cooperative Patent Classification (CPC) codes to measure similarity between firms in the technology space. We classify each patent using 3-digit CPC codes into one of 130 technology classes. Next, for each firm we compute the share of patents in all technology classes and compute the un-centered correlation between firm i's and firm j's patent shares denoted by $\widetilde{\alpha}_{i,j} = \frac{T_i T_j'}{\sqrt{T_i T_i'}\sqrt{T_j T_j'}}$, where T_i is a 130-element vector in which each element corresponds to the number of patents granted to firm i in a given CPC class. Building on Bloom et al. (2013), we account for the firm's importance in the patenting network by weighting each observation by the total citations accrued to the given firm. Let c_j denote all citations attributed to patents assigned to a given firm-window cell. As in the main text, we measure the spillover benefiting a firm i from a firm j as the proximity-weighted, citation-adjusted R&D share $\sigma_{i,j} = \alpha_{i,j}\frac{x_j}{S_i^{ext}}$, where $\alpha_{i,j} = \widetilde{\alpha}_{i,j}\frac{c_j}{C}$, $C = \sum_j c_j$ marks the total citation count, $S_i^{ext} = \sum_{j\neq i}\alpha_{i,j}x_j$ and where innovation rates are recovered from R&D expenses as $x_j \propto s(x_j)^{\frac{1}{\psi}}$.

C.2 BLADE

In what follows, we provide a brief description of the Business Longitudinal Analysis Data Environment (BLADE).

ABS Data Disclaimer. The results of these studies are based, in part, on data supplied to the ABS under the Taxation Administration Act 1953, A New Tax System (Australian Business Number) Act 1999, Australian Border Force Act 2015, Social Security (Administration) Act 1999, A New Tax System (Family Assistance) (Administration) Act 1999, Paid Parental Leave Act 2010 and/or the Student Assistance Act 1973. Such data may only be used for the purpose of administering the Census and Statistics Act 1905 or performance of functions of the ABS as set out in section 6 of the Australian Bureau of Statistics Act 1975. No individual information collected under the Census and Statistics Act 1905 is provided back to custodians for administrative or regulatory purposes. Any discussion of data limitations or weaknesses is in the context of using the data for statistical purposes and is not related to the ability of the data to support the Australian Taxation Office, Australian Business Register, Department of Social Services and/or Department of Home Affairs core operational requirements.

Legislative requirements to ensure privacy and secrecy of these data have been followed. For access to PLIDA and/or BLADE data under Section 16A of the ABS Act 1975 or enabled by section 15 of the Census and Statistics (Information Release and Access) Determination 2018, source data are de-identified and so data about specific individuals has not been viewed in conducting this analysis. In accordance with the Census and Statistics Act 1905, results have been treated where necessary to ensure that they are not likely to enable identification of a particular person or organisation.

Firm-level information. We use administrative firm tax records provided by the Australian Bureau of Statistics as part of the Business Longitudinal Analysis Data Environment (BLADE). BLADE covers the universe of businesses registered for the Goods and Services Tax with total sales exceeding AUD75,000, but does not include sole traders or partnerships that submit personal income tax returns instead of business income tax returns. The subset of BLADE used in this project sources data from the Australian Tax Office for financial years 2005 (July 1 2004-June 30 2005) through to 2018.

Firm-level R&D is reported in the Business Income Tax (BIT) dataset, and is an accounting measure of expenditure subject to the R&D subsidy. R&D expenditure eligible for the subsidy includes salaries and wages, materials and consumables, depreciation of plant and equipment, and overhead expenses. Total sales (turnover) are reported in the Business Activity Statement. We define profits as operating profits, as reported in the BIT.

The ABS classifies firms in BLADE into two groups: the "Non-Profiled Population" and the "Profiled Population." Units in the Non-Profiled Population have a simple structure, and represent the majority of BLADE businesses. The Profiled Population consists of large businesses with complex structures. For businesses in the Profiled Population, the ABS assigns units under common control into an "Enterprise Group." To remain consistent with the Australian Tax Office's assessment of R&D subsidy eligibility, we consolidate accounts in the Profiled Population to the Enterprise Group level.

C.3 Orbis

Orbis is a dataset managed by Bureau van Dijk that collects balance sheet information on both private and public companies across all industrialized countries. As of 2022, the Orbis dataset contains information on around 400 million companies and entities from more than 200 countries and territories, of which more than 99% are private companies. To access the Orbis dataset, we use WRDS services which allows us to retrieve up to 10 years of history of company information.

To retrieve the firm-level data, we filter the database by restricting attention to firm-year observations with positive R&D expenses and to companies registered in one of the 11 countries in the Appelt et al. (2025) dataset.[3] We focus on consolidated balance sheets (Orbis codes "C1", "C2", or "C3"). In calculating the average theoretical elasticity $\epsilon_j = -\frac{1}{\psi-1}\frac{s_j}{\pi_j}$ we follow the same steps as in the case of the Compustat data. Namely, we define profits π_j as sales net of cost of goods sold and R&D expenses. We restrict attention to firms with positive profits, positive revenues and positive cost of goods sold. To account for outliers, we drop observations with $\epsilon_j > 10$. The final sample consists of 16,713 firm-year observations.

D Further Empirical Results and Robustness

D.1 Further Results: Proportionality of Profits and Sales

This Appendix documents that in our Compustat and BLADE samples, profits (and, therefore, their discounted net present values – firm values) are proportional to sales. This provides further support for the theoretical underpinnings of our approach (see Section 2.1).

Formally, we estimate the following regression

$$\log\left(\frac{y_{i,t}}{\pi_{i,t}}\right) = \delta_{i,t} + \alpha t + \beta t^2 + u_{i,t}, \tag{A20}$$

where $\delta_{i,t}$ marks fixed effects (sector, firm, cohort, and their combinations depending on specification) and t indicates time. When conditioning on firm or cohort fixed effects, the regression captures the average evolution of the sales-to-profit ratio over firms' life-cycles.

Table A1 presents the results for Compustat and BLADE. While some coefficients are statistically significant, they are not economically meaningful. For example, after 10

[3] These countries are Australia, Belgium, Czechia, France, Italy, the Netherlands, New Zealand, Norway, Portugal, Slovakia and Sweden.

Table A1: Sales-to-profit ratios over firms' life-cycles: Compustat and BLADE

	Compustat			BLADE		
	(I)	(II)	(III)	(IV)	(V)	(VI)
time	-0.014	-0.003	-0.004	0.008	0.019	0.011
	(0.001)	(0.002)	(0.001)	(0.006)	(0.006)	(0.006)
time^2	0.0002	0.0001	0.0001	0.000	-0.001	0.000
	(0.0001)	(0.0001)	(0.0001)	(0.000)	(0.000)	(0.000)
firm fixed effects	✓			✓		
cohort fixed effects		✓			✓	
sector fixed effects		✓	✓		✓	✓
Observations	125,720	126,852	126,852	98,480	98,480	98,480
R^2	0.58	0.18	0.17	0.66	0.05	0.05
Within R^2	0.008	0.002	0.004	0.001	0.001	0.000

Note: The dependent variable is $\log\left(\frac{y_{i,t}}{\pi_{i,t}}\right)$ which implies that we only consider firms with positive sales and profits. Index t corresponds to a year and, therefore, the "time" variable corresponds to normalized firm age.

years, the sales-to-profit ratio is estimated to decrease (increase) by about 14 (8) *percent* (not percentage points) in Compustat (BLADE).

D.2 Further Results: R&D Cost Elasticity

Our estimates of R&D cost elasticities in the main text focus on average industry-level variation. Specifically, if the estimated γ coefficient is either statistically insignificant or estimated to be above 1–violating our restriction that $\psi = 1/\gamma \geq 1$–we use the unconditional estimates of $\gamma = 0.53$ instead. Furthermore, we shrink the estimated coefficients towards the unconditional estimate by taking a simple average between the unconditional value and estimated coefficients. Both of these changes reduce dispersion in the estimated innovation elasticities. In light of these choices, our results stressing the economic importance of heterogeneity in innovation elasticities should be seen as a conservative lower bound. Estimating cost elasticities at the firm-level results in noisy estimates and, therefore, we use industry-level estimates as our preferred specification.

Next, we investigate whether cost elasticities differ by firm size and age. Specifically, Table A2 demonstrates that our estimates of the R&D cost elasticity are roughly unchanged across the size and age distributions. In particular, while statistically significant, the difference in the R&D elasticity below and above median size (revenues) is quantitatively small. The difference above and below median age is statistically insignificant. Note that the regression includes firm fixed effects which in combination with time fixed effects makes age perfectly collinear. Therefore, in the age regression we include only a linear control for the time window.

Table A2: Size and age-dependent estimates of regression (8)

	(i)	(ii)
R&D expenses	0.54	0.55
	(0.07)	(0.07)
R&D expenses × below median size	0.07	
	(0.03)	
R&D expenses × below median age		0.03
		(0.02)
firm fixed effects	✓	✓
sector fixed effects	✓	✓
time fixed effects	✓	
linear time control		✓
Observations	11,598	11,598

Notes: Poisson regression estimates of regression (8). Outcome variable is the patent count. Data aggregated into non-overlapping 5-year windows. Observations are weighted with log patent citation count. Age corresponds to the number of years since the first occurrence in Compustat. Median size and median age are calculated individually in each time window so that the split regression is not biased towards a specific time period.

D.3 Further Results: BLADE Details

In what follows, we provide further details on the main specification of our BLADE estimation in the main text. In later subsections, we also provide additional robustness checks.

Institutional background. Prior to the 2012 R&D Tax Incentive reform, Australian firms were able to deduct 125% of R&D expenditure from taxable income, as well as an additional 50% of the portion of expenditure exceeding average expenditure over the previous three years.[4] Throughout the sample period, the corporate tax rate was 30% implying an effective R&D subsidy of $1.25 \times 0.3 = 0.375$. Firms with sales less than AUD5 million could instead choose to claim 30% of R&D expenditure as a refundable tax offset i.e. the firm was entitled to a cash refund of any unused offset amount if the firm's tax liability was reduced to zero.

[4] We assume that there was no systematic difference between treatment and control groups with respect to the additional deduction in 2011 i.e. relative to average R&D expenditure over the previous three years, 2011 R&D expenditure was not systematically different between AUD 5-20 million firms and AUD 20+ million firms. The pre-trends shown in Figure 3 provide support for this assumption.

Table A3: Inverse cost elasticity: BLADE

	Patent Applications	Granted Patents
$\gamma = 1/\psi$	0.360	0.396
	(0.098)	(0.135)
firm fixed effects	✓	✓
time fixed effects	✓	✓
Observations	594	368

Note: Estimated coefficient γ in regression (8). Standard errors in parentheses are clustered at the firm level. The coefficient corresponds to the elasticity of patents (innovation output) to R&D expenses (innovation inputs). The sample period is from 2005-2018; to allow for time to build in innovation, we aggregate the data into non-overlapping 5 year windows. The sample consists of Australian firms with average sales between AU$5 million and AU$20 million in the policy reform window (2010-2014) matched to patent data in IPLORD. Each observation corresponds to one firm-window cell. If a firm exists for less than 5 years, we retain the firm in the sample and use all years in which the firm is observed.

Following the policy change in 2012, firms with sales less than AUD20 million were eligible for a refundable 45% tax offset. By contrast, firms with sales exceeding AUD20 million were entitled to a non-refundable 40% tax offset.

Sample selection. In keeping with the Compustat analysis, we consider firms that report positive R&D expenditure and positive profits in both 2011 and 2012. To deal with outliers, we drop the top 5% of firms with regard to their R&D-to-profit ratio.

Further details on estimation methodology. For the purposes of the difference-in-differences estimation, we only consider firms with sales exceeding AUD5 million.[5] In addition, we only consider firms that remain in the same sales category (i.e. AUD5-20 million or AUD20+ million) across both 2011 and 2012 (the vast majority of firms). Table A7 shows that the findings are largely unchanged if we include firms that switch sales categories and fix the treatment and control groups based on pre-treatment status i.e. 2011 sales.

To convert the point estimate of β_1 from (9) to an R&D elasticity, we divide by -0.087, which represents the log difference in the change of subsidy between the treatment and control groups: $\Delta \log P_s = \log((1-0.45)/(1-0.375)) - \log((1-0.4)/(1-0.375)) = -0.087$.

[5]Depending on firm profitability and R&D expenditure, firms with sales less than AUD5 million in 2011 may have been better off claiming the refundable tax offset (30%) rather than the non-refundable tax deduction (37.5%) or vice versa, complicating the calculation of firm-level responsiveness to the subsequent subsidy change.

Table A4: Asymptotic approximation standard errors for the relative Bang for the Buck.

	A_k/A	Lower bound of 95% CI	Upper bound of 95% CI
R&D-int.	0.711	0.707	0.715
Young	1.376	1.375	1.377
Small	1.329	1.329	1.330
- Young	1.524	1.523	1.525
- Old	0.807	0.802	0.812
High-gr.	1.441	1.440	1.443
Random	1.000	0.999	1.001

Notes: The table presents relative Bangs for the Buck and the asymptotic confidence intervals of the form $\pm 1.96\sqrt{\text{var}(A_k)}$ where $\text{var}(A_k)$ is defined in (A21).

D.4 Further Results: Standard Errors of Relative Bang for the Buck

In this section we quantify the statistical significance of the estimated Bangs for the Buck and the difference in relative cost-effectiveness of the targeted subsidies between groups of firms. Let the population value of the Bang and the Buck be denoted by $\overline{B}_k = \sum_{j\in\overline{\Omega}_k} b_j$ and $\overline{C}_k = \sum_{j\in\overline{\Omega}_k} c_j$, respectively, where $\overline{\Omega}_k$ marks the theoretical population of firms of type k from which the Compustat data is sampled. Assume that firm level components are normally distributed. Then, the joint distribution is

$$\frac{1}{N_k}\begin{pmatrix} B_k - \overline{B}_k \\ C_k - \overline{C}_k \end{pmatrix} \to N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{B,k} & \sigma_{B,C,k} \\ \sigma_{B,C,k} & \sigma^2_{C,k} \end{pmatrix}\right)$$

Consider now the ratio $\frac{B_k}{C_k} = \frac{B_k^* + \overline{B}_k}{C_k^* + \overline{C}_k}$ where B_k^* and C_k^* are zero mean Gaussian random variables.

Applying the delta method, we obtain the expression for the asymptotic variance of the ratio $A_k = \frac{B_k}{C_k}$

$$\text{var}(A_k) = \frac{1}{\overline{C}_k^2}\left(\sigma^2_{B,k} + \frac{\overline{B}_k^2}{\overline{C}_k^2}\sigma^2_{C,k} - 2\frac{\overline{B}_k}{\overline{C}_k}\sigma_{B,C,k}\right). \tag{A21}$$

Table A4 presents standard errors of the Bangs for the Buck of the form $\pm\alpha\sqrt{\widehat{\text{var}}(A_k)}$, where $\widehat{\text{var}}(A_k)$ is the consistent estimate of $\text{var}(A_k)$ defined in (A21), in which population moments are replaced by sample analogs and where α corresponds to a percentile of the normal distribution. The implied errors are very tight and all differences between groups are statistically significant.

However, the resulting confidence intervals may become distorted in finite samples.

Table A5: Bootstrap standard errors for the relative Bangs for the Buck.

	A_k/A	Lower bound of 95% CI	Upper bound of 95% CI
R&D-int.	0.71	0.61	0.81
Young	1.38	1.25	1.51
Small	1.33	1.20	1.48
- Young	1.52	1.37	1.69
- Old	0.81	0.63	1.04
High-gr.	1.44	1.31	1.59
Random	1.00	1.00	1.00

Notes: The table presents relative Bangs for the Buck and the standard errors of the form $\left[q_{2.5}^{A_k}, q_{97.5}^{A_k}\right]$, where $q_x^{A_k}$ is the x-th percentile of the bootstrap distribution of the Bangs for the Buck A_k. We use 1000 bootstrap samples.

To see why, note that we can write $\frac{B_k}{C_k} = \frac{\overline{B}_k}{\overline{C}_k}\frac{1+B_k^*/\overline{B}_k}{1+C_k^*/\overline{C}_k}$. If the population value of the cost $\overline{C}_k \geq 0$ in some subgroup k is quantitatively close to zero and therefore highly skewed, the normal approximation is inaccurate. This is precisely the case in our application where individual values of cost are non-negative, and the sum may be arbitrarily close to zero. To address potential inaccuracy of the asymptotic first-order approximation, below we provide bootstrap standard errors.

Bootstrap standard errors. To estimate standard errors for the relative bangs for the buck, we fix the assignment of firms to groups based on the original sample. Next, we draw observations with replacement from the original Compustat sample to obtain a bootstrap sample of the same size.[6] In each of the 1000 bootstrap samples, we repeat the procedure of computing the relative bangs for the buck as described in the main text.

In Table A5, we report the confidence interval based on the percentiles of the distribution of bootstrap Bangs for the Buck. While the confidence intervals are much wider than their asymptotic counterparts, our main results remain statistically significant. Small-young firms are a statistically significantly more cost-effective target than small-old firms. R&D intensive firms are statistically speaking as inefficient a target for R&D subsidies as small-old firms. High-growth, and young and small-young firms are the most cost-efficient to support.

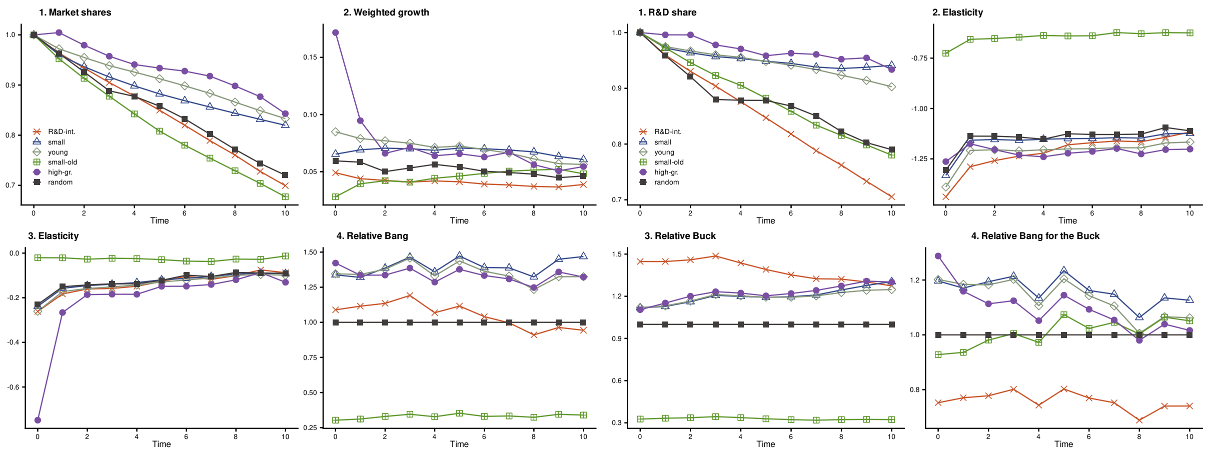
D.5 Further Results: Dynamics of the Bangs for the Buck

In this section we present more details pertaining to the dynamics of the estimated Bangs for the Buck across our firm groups. Our goal is to estimate how the relative Bang and

[6]We keep the assignment to groups fixed because our assignment method chooses firms with, say, highest growth rate up to 50% market share. Therefore, the assignment method can to a large extent reverse the impact of randomization.

Figure A1: Persistence of the Bang for the Buck and its components using unbalanced panel

(a) Drivers of the Bang (b) Drivers of the Buck



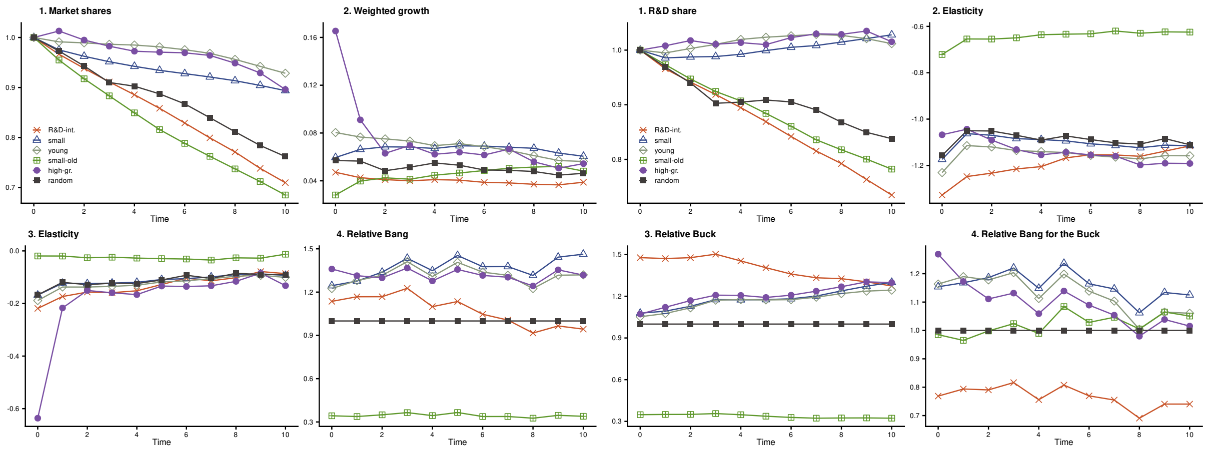
Note: Figure presents components of the Bang and the Buck using annual data and an unbalanced panel of firms. See also notes to Figure 5.

Buck change over time for fixed firm groups. In other words, is supporting firms that were high-growth one year or two years ago as sound a policy as supporting the currently high-growth ones? To answer this question, we pool all firms and classify them into groups in the same manner as in Section 3.4. Then, keeping this classification fixed, we track firms over time and estimate all the components of the Buck and the Bang. For example, we estimate the efficiency of supporting firms that were high-growth one year ago, two years ago, etc. Or, in the case of window-averaged data, in the previous window, two windows ago, etc. As reported in the main text, young firms tend to exhibit a persistently high benefits-to-cost ratio and stand out as excellent targets for R&D subsidies. High-growth firms, on the other hand, while the most cost-effective target from an instantaneous perspective, exhibit a rapid drop over time in their relative Bang for the Buck. To understand the forces behind the patterns in the Bangs for the Buck, in this section we present the evolution of the components of this aggregate measure.

Dynamics of Bang for the Buck Components. Figure A1 presents the evolution of the components of the Bang and the Buck over time in an unbalanced panel of firms. Panel (a) reports the components of the Bang. Sub-panel 1. shows the market share m across firm groups, normalized to 1 at time $t = 0$. The panel illustrates that small and young firms initially gain market share, but all firms converge to a similar size in the long run. Old firms tend to shrink on average, as illustrated by a sharp decline in the market share of all firms and small-old businesses. This result is consistent with evidence on the universe of firms in the US, reported in Haltiwanger et al. (2013): the net growth in the economy is entirely due to young firms.

Figure A2: Persistence of the Bang for the Buck and its components using balanced panel

(a) Drivers of the Bang (b) Drivers of the Buck



Note: Figure presents components of the Bang and the Buck using annual data and a balanced panel of firms. See also notes to Figure 5.

Next in sub-panel 2., we can see that the size-weighted growth of high-growth firms drops quickly over time. In particular, firms that were classified as high-growth 2 years ago grow at a slower rate than those that were small or young two years ago. Small firms exhibit the highest growth rates from the long run perspective. However, this is driven entirely by the fact that most of them are young. This can be seen from the poor growth trajectory of small-old firms. Sub-panel 3. illustrates convergence in Bang-relevant elasticities. Note that after about the first 2 years, these elasticities remain relatively stable, consistent with our baseline results. Finally, sub-panel 4. concludes by plotting the relative Bangs over time.

Next, Panel (b) on the right-hand side reports the components of the Buck and the total relative Bangs for the Buck in sub-panel 4. Note that sub-panel 4. replicates Panel (a) in Figure 5. Here, the most striking result is the stable R&D share among high-growth firms. Recall from our previous discussion that their revenue growth and market shares are gradually declining, reducing the Bang. In combination with their stable R&D share, this is responsible for the decline in their relative Bang for the Buck visible in sub-panel 4.

Figure A2 presents the same set of results as above, but generated from the balanced panel of firms. That is, we restrict attention to the set of firms for which we can obtain all necessary balance sheet items in at least 10 consecutive years. Here sub-panel 4. in Panel (b) corresponds to the results reported in Figure 5 in the main text. By focusing on a balanced panel, we select firms based on their ex-post success. This is particularly visible when inspecting the time paths for young firms: ex-post successful young firms grow rapidly, increasing their market and R&D shares, as can be seen in sub-panels 1. in Panels (a) and (b) in Figure A2. This elevates their relative Bang for the Buck throughout the horizon of interest. The dynamics of the remaining firm groups are broadly in line with the results in the unbalanced panel described above.

D.6 Robustness: BLADE Fuzzy Difference-in-Differences

Our baseline difference-in-differences regression (9) estimates the intention-to-treat (ITT) effect, i.e. the responsiveness of firms to the policy reform based on subsidy *eligibility*. This is irrespective of whether the firm actually took up the more generous subsidy. In this section, we consider a sub-sample of firms for which we *observe the choice* of subsidy.

We find that some firms in the AU\$5 − 20 million category did not choose the more generous 45% offset, instead opting for the 40% offset.[7] We, therefore, employ a "fuzzy" DiD design (the fuzziness arising from the fact that some units in the treatment group do not take the treatment), using the Wald-DiD estimator.

In addition to the standard parallel trends assumption, de Chaisemartin and D'Haultfoeuille (2018) highlight further assumptions required for the Wald-DiD to estimate the local average treatment effect (LATE). For our two-period setting, in which all units are untreated in the pre-period (i.e. 2011), we require that either (1) there is a stable percentage of treated units in the control group, or (2) the LATE is homogeneous across both treatment and control groups. Since the control group is ineligible for the 45% offset, assumption (1) is satisfied.

As such, we run an instrumental variable regression of R&D expenditure on the 45% offset, with year and sales category as included instruments, and the interaction of the two as the excluded instrument. In particular, the first stage is:

$$\mathbb{1}_{45\%} = \kappa_0 + \kappa_1 \mathbb{1}_{\$5-20M} \times \mathbb{1}_{2012} + \kappa_2 \mathbb{1}_{\$5-20M} + \kappa_3 \mathbb{1}_{2012} + \kappa_4 X_{i,t} + \eta_{i,t}, \tag{A22}$$

where $\mathbb{1}_{45\%}$ is an indicator function equal to one if firm i took the 45% offset and all other variables are the same as in the main text.

From regression (A22), we recover the predicted value, $\widehat{\mathbb{1}}_{45\%}$, and estimate the second stage:

$$\log(R\&D)_{i,t} = \alpha + \theta_1 \widehat{\mathbb{1}}_{45\%} + \gamma_1 \mathbb{1}_{\$5-20M} + \delta_1 \mathbb{1}_{2012} + \lambda X_{i,t} + u_{i,t}, \tag{A23}$$

where the coefficient θ_1 is the LATE among compliers i.e. AUD5-20 million firms that took the 45% offset.

[7] Survey evidence suggests that some firms remain unaware of R&D subsidies for many years after such policies are introduced (Thomson and Webster, 2012).

Table A6: Fuzzy Difference-in-Differences: BLADE

	(I)	(II)	(III)	(IV)
45% Subsidy, θ_1	0.202	0.208	0.287	0.303
	(0.084)	(0.084)	(0.092)	(0.093)
45% Subsidy $\times$ e, θ_2			-0.092	-0.102
			(0.056)	(0.058)
	Theoretical Prediction			
$\theta_1 = -\frac{1}{N}\sum_i \frac{s_i}{\pi_i}\frac{\psi}{\psi-1}\Delta \log P_s$	0.144	0.144		
$\theta_2 = 1 \times \Delta \log P_s$			-0.087	-0.087
Control for sales		✓		✓
Observations	1502	1502	1502	1502

Notes: The table reports the regression results from (A23) and (A26). Columns refer to different specifications (with and without sales and with and without R&D to profits as a control). The regression sample includes firms in years 2011 and 2012 with sales exceeding AU$5 million, those that report positive R&D expenditure, positive operating profits and remain in the same sales category across both 2011 and 2012, and for which the choice of R&D subsidy is observed. Standard errors in parentheses are clustered at the firm level.

In the heterogeneous specification, we require two first stages:

$$\begin{aligned}\mathbb{1}_{45\%} &= \kappa_0 + (\kappa_1 + \kappa_2 e_i)\mathbb{1}_{\$5-20M} \times \mathbb{1}_{2012} \\ &\quad + (\kappa_3 + \kappa_4 e_i)\mathbb{1}_{\$5-20M} + (\kappa_5 + \kappa_6 e_i)\mathbb{1}_{2012} + \kappa_7 e_i + \kappa_8 X_{i,t} + \eta_{i,t}, \qquad \text{(A24)} \\ \mathbb{1}_{45\%} \times e_i &= \phi_0 + (\phi_1 + \phi_2 e_i)\mathbb{1}_{\$5-20M} \times \mathbb{1}_{2012} \\ &\quad + (\phi_3 + \phi_4 e_i)\mathbb{1}_{\$5-20M} + (\phi_5 + \phi_6 e_i)\mathbb{1}_{2012} + \phi_7 e_i + \phi_8 X_{i,t} + \varepsilon_{i,t}. \qquad \text{(A25)}\end{aligned}$$

Recovering $\widehat{\mathbb{1}}_{45\%}$ and $\mathbb{1}\widehat{_{45\%} \times e_i}$, we estimate the second stage:

$$\begin{aligned}\log(R\&D)_{i,t} =& \alpha + \theta_1 \widehat{\mathbb{1}}_{45\%} \times \mathbb{1}_{2012} + \theta_2 \mathbb{1}\widehat{_{45\%} \times e_i} \times \mathbb{1}_{2012} \\ &+ \gamma_1 \widehat{\mathbb{1}}_{45\%} + \gamma_2 \mathbb{1}\widehat{_{45\%} \times e_i} + \delta_1 \mathbb{1}_{2012} + \delta_2 e_i \times \mathbb{1}_{2012} + \eta e_i + \lambda X_{i,t} + u_{i,t}. \qquad \text{(A26)}\end{aligned}$$

Table A6 presents the results from the Wald-DiD. First, note that the estimates of θ_1 (the LATE) are greater than those of β_1 (the ITT) in Table 3. This reflects the fact that not all firms took advantage of the policy. As in the main text, using Proposition 1, we can write $\theta_1 = \Delta \log P_s \bar{\epsilon}^s = \Delta \log P_s \frac{1}{N}\sum_i -\frac{s_i}{\pi_i}\frac{\psi}{\psi-1}$. For the sample of complier firms, this value is about $\Delta \log P_s \bar{\epsilon}^s = 0.144$, so we are unable to reject the null hypothesis in the first two columns of Table A6 (p-values of 0.497 and 0.449, respectively). In columns 3 and 4 of Table A6 we also cannot reject $H_0 : \theta_2 = 1 \times \Delta \log P_s$ (p-values of 0.923 and 0.802, respectively).

Table A7: Sales Category Switchers: BLADE

	(I)	(II)	(III)	(IV)
AU\$5 − 20 mil.$_{2011}$ × 2012, β_1	0.123	0.152	0.184	0.219
	(0.054)	(0.053)	(0.055)	(0.053)
AU\$5 − 20 mil.$_{2011}$ × 2012 × e, β_2			-0.121	-0.122
			(0.033)	(0.033)
	Theoretical Prediction			
$\beta_1 = -\frac{1}{N}\sum_i \frac{s_i}{\pi_i}\frac{\psi}{\psi-1}\Delta \log P_s$	0.129	0.129		
$\beta_2 = 1 \times \Delta \log P_s$			-0.087	-0.087
Control for sales		✓		✓
Observations	1934	1934	1934	1934

Notes: The table reports the regression results from (9) and (10), where AU\$5 − 20 mil$_{2011}$ is an indicator variable equal to one for firms that report sales between AU\$5 million and AU\$20 million in 2011, and zero for firms that report sales exceeding AU\$20 million in 2011. Columns refer to different specifications (with and without sales and with and without R&D to profits as a control). The regression sample includes firms in years 2011 and 2012 that report positive R&D expenditure and positive operating profits, and sales exceeding AU\$5 million in 2011. Standard errors in parentheses are clustered at the firm level.

D.7 Robustness: BLADE Sales Category Switchers

In our baseline specification of Table 3, we exclude firms that change sales categories between 2011 and 2012 i.e. firms that report sales between AU\$5 million and AU\$20 million in 2011 and sales greater than AU\$20 million in 2012, or vice versa. If these "switcher" firms were able to manipulate sales to receive the more generous subsidy, defining treatment status based on post-reform sales in 2012 may introduce selection bias. For instance, if firms that were planning to increase R&D expenditures in 2012 were also more likely to switch into the treatment group, this would undermine the validity of the parallel trends assumption.

As a robustness check, we instead fix the treatment and control groups based on sales in 2011 (the pre-treatment period). This approach estimates the effect of offering the 45% subsidy to firms that were eligible prior to the policy reform. Table A7 presents the regression results. As in the main text, the first two columns show β_1 estimated with and without including sales as a control. We are unable to reject $H_0 : \beta_1 = \Delta \log P_s \bar{\epsilon}^s$ (p-values of 0.921 and 0.654, respectively). In the final two columns we also cannot reject $H_0 : \beta_2 = \Delta \log P_s$ (p-values of 0.314 and 0.293, respectively).

We now consider size category switchers in the context of our fuzzy difference-in-differences design. In particular, we run an instrumental variable regression of R&D expenditure on the 45% subsidy, with year and sales category (defined based on 2011 sales) as included instruments, and the interaction of the two as the excluded instrument. The estimate of θ_1 from regression (A23) is the local average treatment effect among complier

Table A8: Fuzzy Difference-in-Differences with Sales Category Switchers: BLADE

	(I)	(II)	(III)	(IV)
45% Subsidy, θ_1	0.154	0.193	0.232	0.290
	(0.088)	(0.085)	(0.096)	(0.092)
45% Subsidy $\times$ e, θ_2			-0.130	-0.131
			(0.054)	(0.054)
	Theoretical Prediction			
$\theta_1 = -\frac{1}{N}\sum_i \frac{s_i}{\pi_i}\frac{\psi}{\psi-1}\Delta \log P_s$	0.148	0.148		
$\theta_2 = 1 \times \Delta \log P_s$			-0.087	-0.087
Control for sales		✓		✓
Observations	1628	1628	1628	1628

Notes: The table reports the regression results from (A23) and (A26). Columns refer to different specifications (with and without sales and with and without R&D to profits as a control). The regression sample includes firms in years 2011 and 2012 with sales exceeding AU$5 million, those that report positive R&D expenditure, positive operating profits, sales exceeding AU$5 million in 2011 and for which the choice of R&D subsidy is observed. Standard errors in parentheses are clustered at the firm level.

firms. In this case, compliers are firms that were both eligible for the 45% subsidy prior to the reform and ultimately claimed the 45% subsidy.

Table A8 presents the results from the Wald-DiD with sales category switchers. Again, we are unable to reject $H_0 : \theta_1 = \Delta \log P_s \bar{\epsilon}^s$ in the first two columns (p-values of 0.940 and 0.593, respectively). In the last two columns we also cannot reject $H_0 : \theta_2 = 1 \times \Delta \log P_s$ (p-values of 0.431 and 0.421, respectively).

D.8 Robustness: Discounting

The exposition in this section, as well as our methodological approach, closely follows Gormsen and Huber (2025). The Weighted Average Cost of Capital (WACC) represents a firm's average cost of capital from all sources, including common stock, preferred stock, bonds, and any other long-term debt. It is designed to capture the market value of a firm's cash flows while adjusting for leverage in the form of long-term debt. WACC consists of the cost of equity and the cost of debt, weighted by the leverage ratio (i.e., the value of debt relative to the total value of capital). These two components reflect the value of the investment decision and the financing decision, respectively (Miles and Ezzell, 1980).

Let $v = e + d$ denote the firm value, equal to the sum of equity e and debt d. The respective marginal cost of raising an additional unit of capital is R_e for equity and R_d for debt.

$$WACC = \frac{e}{v} \times R_e + \frac{d}{v} \times R_d \times (1 - T) \tag{A27}$$

where $(1-T)$ adjusts the cost of debt for the fact that interest payments are tax-deductible.

We assume a corporate income tax rate of $T = 0.21$. The equity value e corresponds to the market capitalization, i.e., share price multiplied by the number of shares outstanding (Compustat mnemonics `prcc_f` and `csho`, respectively). We let debt value d equal the book value of long-term debt (maturity above 1 year, `dltt`), but our results are robust to the inclusion of current debt as well. We discount future profits and R&D expenditures using the WACC corrected for aggregate nominal economic growth (g), as in the Gordon Growth Model.

Cost of Debt. To calculate the cost of debt for each firm, we use Standard & Poors (S&P) credit ratings as a proxy for credit access. We utilize the WRDS "adsprate" database, which contains daily S&P rating updates spanning the years 1981 to 2017. S&P scores firms on a 23-point scale ranging from "AAA" (the most creditworthy) to "BBB-" (the lowest "investment grade" rating), "CCC" (vulnerable), and "SD" or "D" for selective default and default, respectively.

We obtain ratings for 5,972 unique firms out of our total sample of 17,825. Ratings are updated daily, and we aggregate them within years or averaging windows. To facilitate this, we assign numerical values to the ratings—ranging from 1 for AAA to 23 for SD—and use the median value within the specific year or window.

If no rating updates occur within a given cell, we use the most recent available rating. If ratings are missing at the beginning of the sample for a given firm, we back-fill using the first available rating.

Given the median rating for a period, we assign the cost of debt based on Moody's Seasoned Corporate Bond Yield Index, corresponding to AAA, BAA, and High Yield (i.e., below investment grade) bonds.[8] For median numerical ratings between AAA and AA-, we assign the yield of the AAA index. For the remaining investment-grade ratings (A+ to BBB-), we assign the BAA yield. For firms with ratings below BBB-, we assign the interest rate indicated by the High Yield Index. For firms missing an S&P credit rating, we assign a cost of debt equal to the BAA yield.

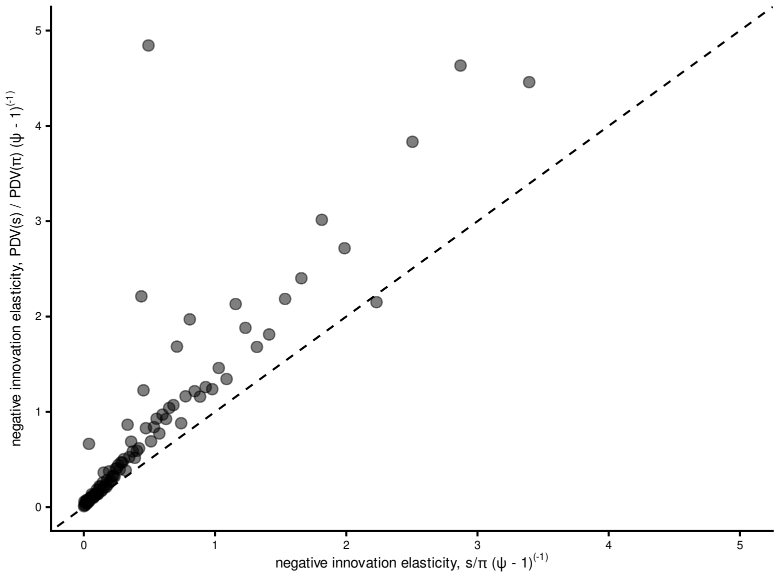
Cost of Equity. The cost of equity captures the opportunity cost of a firm's capital stock (Modigliani and Miller, 1963). To estimate this, we employ the Capital Asset Pricing Model (CAPM). According to this model, the expected return on equity is expressed as:

$$E[R_e] = R_f + \beta \varrho, \tag{A28}$$

where R_f is the risk-free interest rate and $\varrho \equiv R_m - R_f$ represents the equity risk premium—the difference between the market return (R_m) and the risk-free rate (R_f).

[8] FRED database tickers: `DAAA` for AAA yield, `DBAA` for BAA yield, and `BAMLH0A0HYM2EY` for high yield bonds.

Figure A3: Innovation elasticities: PDV with WACC and special cases.



The figure presents a binned scatter plot of the "special case" innovation elasticities – equation (7) – and the "PDV-based" innovation elasticities – equation (6). Each dot represents 1 percent of the firm-window observations in our sample. The dashed line is the 45 degree line. The discount rate is computed as the weighted average cost of capital (see Section D.8). To account for outliers, we restrict attention to firms with elasticities below 10 (in either of the two ways of computing elasticities). The median number of 5-year periods over which we calculate the PDV is four.

We assume the risk-free rate corresponds to the yield on 10-year US Treasuries. The coefficient β is a firm-specific measure capturing the sensitivity of the firm's stock returns to market returns. For the market return R_m, we use the returns of the S&P 500 index.

We estimate firm-specific β coefficients across non-overlapping five-year windows. We merge our primary Compustat sample with CRSP monthly stock return data for individual stocks and the market index. To obtain β, we estimate the following regression:

$$R_{i,t,w} = \alpha_{i,w} + \beta^e_{i,w} R_{m,t,w} \tag{A29}$$

where i indexes firms, t months, and w non-overlapping five-year windows. Following Frazzini and Pedersen (2014), we account for noise in the estimates by winsorizing the top and bottom 1% of coefficients and shrinking the remaining betas toward their unconditional mean. Formally, we calculate:

$$\widehat{\beta}_{i,w} = (1-\xi)\beta^e_{i,w} + \xi\overline{\beta},$$

where, following Gormsen and Huber (2025), we set $\xi = 0.4$ and $\overline{\beta} = 1$. We back-fill and forward-fill missing β coefficients where necessary. If we are unable to estimate β for a specific firm-window cell, we assume $\beta = 1$.

Table A9: Decomposition of the Bang for the Buck for employment-based growth

	Relative Bang & Buck			Bang Components				Buck Components		
Firm group	A_k/A	B_k/B	C_k/C	m_k	$\bar{\epsilon}_k^B$	g_k	θ_k^g	r_k	$\bar{\epsilon}_k^s$	θ_k^s
R&D-int.	0.66	0.94	1.43	0.50	-0.23	0.02	6.22	0.73	-1.34	0.92
Young	1.27	1.44	1.13	0.50	-0.34	0.07	2.05	0.55	-1.45	1.13
Small	1.26	1.42	1.12	0.50	-0.31	0.08	1.76	0.55	-1.38	1.22
- Young	1.36	1.11	0.82	0.34	-0.34	0.07	2.16	0.39	-1.45	1.30
- Old	1.00	0.30	0.30	0.16	-0.01	0.11	26.42	0.16	-0.62	1.83
High-gr.	1.26	1.53	1.21	0.50	-0.81	0.14	0.44	0.61	-1.37	1.01
Random	1.00	1.00	1.00	0.51	-0.31	0.06	1.68	0.51	-1.34	0.93

Note: The table reproduces Table 5 when using employment growth, rather than sales, as the measure of firm growth. See also the notes to Table 5. High-growth firms are defined as firms with *employment growth* above 20%. The market share of "Random" may deviate slightly from 50% due to a finite sample.

In summary, the estimated WACC is expressed as:

$$\widehat{WACC}_{i,w} = \frac{e_{i,w}}{v_{i,w}} \times \left(R_{f,w} + \widehat{\beta}_{i,w}\widehat{\varrho}_w\right) + \frac{d_{i,w}}{v_{i,w}} \times R(\text{rating}_{i,w}) \times (1 - 0.21)$$

where $R_{f,w}$ is the average 10-year Treasury rate in window w. For annual data, w corresponds to a single year; for smoothed data, w represents the five-year window.[9]

Finally, to discount future cash flows and R&D expenses, we use $\widehat{WACC}_{i,w} - g$, where $g = 0.052$ is the average growth rate of nominal GDP in our sample period. We cap the WACC at 25% and drop observations with implied negative WACC.

Figure A3 documents that results are materially unchanged.

D.9 Robustness: Definition of Firm-Level Growth

Table A9 shows that if firm-level growth rate g_j is measured using employment, rather than revenues, the results remain quantitatively and qualitatively unchanged. Note that we have defined high-growth firms as firms with high revenue growth, which does not automatically imply rapid employment growth. Nevertheless, high-growth firms report the highest average employment growth of the considered firm groups. Similarly, small firms remain defined as the firms with below-median sales. All in all, these results further strengthen our main message that high-growth and young firms are the most cost-effective firm group to support.

[9]To minimize noise, market β estimates are always calculated over five-year periods.

Table A10: Decomposition of the Bang for the Buck with alternative profit measure.

	Relative Bang & Buck			Bang Components				Buck Components		
Firm group	A_k/A	B_k/B	C_k/C	m_k	$\bar{\epsilon}_k^B$	g_k	θ_k^g	r_k	$\bar{\epsilon}_k^s$	θ_k^s
R&D-int.	0.61	0.87	1.43	0.50	-0.32	0.04	2.37	0.73	-1.34	0.91
Young	1.20	1.36	1.13	0.50	-0.41	0.08	1.54	0.55	-1.46	1.12
Small	1.15	1.30	1.13	0.50	-0.38	0.06	2.13	0.55	-1.39	1.21
- Young	1.32	1.09	0.83	0.34	-0.41	0.08	1.87	0.39	-1.47	1.29
- Old	0.70	0.21	0.30	0.16	-0.04	0.02	29.49	0.15	-0.62	1.83
High-gr.	1.26	1.45	1.15	0.50	-0.82	0.13	0.51	0.58	-1.38	1.00
Random	1.00	1.00	1.00	0.50	-0.36	0.05	1.90	0.50	-1.35	1.00

Note: The results in the table are based on the profits defined as the income before interest and depreciation. The sample of firms differs from the one underlying Table 5 due to missing profit data. See also the notes to Table 5. The market share of "Random" may deviate slightly from 50% due to a finite sample.

D.10 Robustness: Definition of Profits

Table A10 indicates that our results are robust to an alternative definition of profits. In the main text, we use a model-consistent measure defined as revenues, less costs of goods sold and R&D expenses. A broader profitability measure, operating income before depreciation and amortization (Compustat mnemonic `oibdp`) includes other expenses, such as marketing. When using this measure of profitability in the definition of elasticity ϵ, all results remain qualitatively and quantitatively unchanged.

D.11 Robustness: Time Averaging

The goal of this section is to document that the results obtained in the main text are not an artifact of the averaging procedure that we employ. Table A11 presents the results using annual data rather than averaging over non-overlapping 5-year windows. Each firm-year cell is treated as an independent observation. The results in Table A11 are very similar to those in Table 5 in the main text.

D.12 Robustness: Non-negative Growth Rates

Table A12 documents that the relative ranking of targeted subsidies remains intact when we restrict attention to the firms with positive revenue growth rates.

D.13 Robustness: Negative Profits

In the main text, we consider only firms with positive profits. The reason is that the environment used to derive the theoretical R&D elasticity ϵ implicitly assumes that all firms record positive profits. In this Appendix, we document that our results are robust to

Table A11: Decomposition of the Bang for the Buck using annual data.

Firm group	Relative Bang & Buck A_k/A	B_k/B	C_k/C	Bang Components m_k	$\bar{\epsilon}_k^B$	g_k	θ_k^g	Buck Components r_k	$\bar{\epsilon}_k^s$	θ_k^s
R&D-int.	0.75	1.09	1.45	0.50	-0.26	0.05	4.11	0.73	-1.44	0.93
Young	1.20	1.35	1.12	0.50	-0.26	0.08	2.94	0.55	-1.39	1.19
Small	1.20	1.34	1.12	0.50	-0.24	0.07	4.16	0.55	-1.33	1.27
- Young	1.31	1.04	0.79	0.33	-0.26	0.08	3.42	0.38	-1.40	1.35
- Old	0.93	0.30	0.33	0.17	-0.02	0.03	75.82	0.17	-0.73	1.72
High-gr.	1.29	1.42	1.10	0.50	-0.75	0.17	0.54	0.55	-1.26	1.12
Random	1.00	1.00	1.00	0.49	-0.23	0.06	3.62	0.51	-1.31	1.00

Note: The table reproduces Table 5 when we use annual data instead of non-overlapping 5-year averages. See also the notes to Table 5. The market share of "Random" may deviate slightly from 50% due to a finite sample.

Table A12: Decomposition of the Bang for the Buck for $g_j > 0$

Firm group	Relative Bang & Buck A_k/A	B_k/B	C_k/C	Bang Components m_k	$\bar{\epsilon}_k^B$	g_k	θ_k^g	Buck Components r_k	$\bar{\epsilon}_k^s$	θ_k^s
R&D-int.	0.67	0.96	1.43	0.50	-0.75	0.09	0.61	0.72	-1.30	0.93
Young	1.32	1.55	1.18	0.50	-0.83	0.13	0.58	0.57	-1.40	1.15
Small	1.24	1.42	1.14	0.50	-0.79	0.11	0.66	0.55	-1.34	1.24
- Young	1.49	1.23	0.82	0.34	-0.84	0.13	0.69	0.39	-1.41	1.32
- Old	0.60	0.19	0.32	0.16	-0.39	0.07	0.88	0.16	-0.66	1.74
High-gr.	1.46	1.82	1.25	0.50	-0.85	0.17	0.52	0.61	-1.43	1.03
Random	1.00	1.00	1.00	0.51	-0.78	0.09	0.57	0.50	-1.31	0.99

Note: The table reproduces Table 5 when we restrict attention to firms with positive growth rates. See also the notes to Table 5.The market share of "Random" may deviate slightly from 50% due to a finite sample.

two ways of dealing with firms reporting negative profits. First, by adjusting the definition of the innovation elasticity. Second, by using firm value instead.

Including firms with negative profits. If firms report negative profits, their innovation elasticity would imply a reduction in R&D expenses following a lowering of the price of innovation. To deal with this, we impose that the innovation elasticity for firms with negative profits is given by $\epsilon = \frac{1}{\psi-1}\frac{s}{\pi} < 0$. In other words, we treat the R&D-to-profit ratio as informative about the R&D intensity of firms, but we make sure that the elasticity is always negative. Table A13 shows that our results hold in the extended sample.

Table A13: Decomposition of the Bang for the Buck including negative profits

	Relative Bang & Buck			Bang Components				Buck Components		
Firm group	A_k/A	B_k/B	C_k/C	m_k	$\bar{\epsilon}_k^B$	g_k	θ_k^g	r_k	$\bar{\epsilon}_k^s$	θ_k^s
R&D-int.	0.67	0.88	1.31	0.50	-0.33	0.04	2.61	0.71	-1.46	0.90
Young	1.36	1.51	1.11	0.50	-0.37	0.08	2.04	0.57	-1.69	1.10
Small	1.32	1.46	1.10	0.50	-0.34	0.06	2.74	0.56	-1.62	1.18
- Young	1.52	1.26	0.83	0.34	-0.37	0.08	2.48	0.41	-1.70	1.29
- Old	0.72	0.20	0.28	0.16	-0.02	0.02	52.91	0.15	-0.68	1.68
High-gr.	1.37	1.52	1.11	0.50	-0.93	0.13	0.50	0.59	-1.58	0.96
Random	1.00	1.00	1.00	0.53	-0.34	0.05	2.33	0.53	-1.61	0.89

Note: The table reproduces Table 5 when we include firms reporting negative profits. We use an ad-hoc assumption that the elasticity is proportional to the absolute value of the R&D-to-profit ratio. See also the notes to Table 5. The market share of "Random" may deviate slightly from 50% due to a finite sample.

Table A14: Decomposition of the Bang for the Buck including negative profits, alternative elasticity measure

	Relative Bang & Buck			Bang Components				Buck Components		
Firm group	A_k/A	B_k/B	C_k/C	m_k	$\bar{\epsilon}_k^B$	g_k	θ_k^g	r_k	$\bar{\epsilon}_k^s$	θ_k^s
R&D-int.	0.71	0.93	1.30	0.50	-0.70	0.04	2.80	0.69	-3.35	1.03
Young	1.18	1.28	1.09	0.50	-0.65	0.08	2.20	0.58	-3.23	1.06
Small	1.20	1.30	1.08	0.50	-0.61	0.07	3.08	0.57	-3.15	1.13
- Young	1.26	1.04	0.82	0.34	-0.65	0.08	2.60	0.43	-3.24	1.17
- Old	1.00	0.26	0.26	0.16	-0.05	0.02	61.71	0.15	-1.89	1.50
High-gr.	1.33	1.43	1.08	0.50	-1.65	0.14	0.60	0.59	-2.85	1.07
Random	1.00	1.00	1.00	0.52	-0.63	0.05	2.78	0.53	-3.09	1.11

Note: The table reproduces Table 5 when we include firms reporting negative profits. To calculate the elasticity, we use formula (A5). In the formula, we let $\sum_j \beta^j s_{t+j} = \frac{s_t}{1-\beta}$ where $\beta = 0.96$. The firm value v is measured as market value. See also the notes to Table 5. The market share of "Random" may deviate slightly from 50% due to a finite sample.

Using firm value. We measure v_t as the company-level consolidated market value.[10] As for the present discounted value of future R&D expenses, we assume that the firm expects them to remain constant over time, $\sum_j \beta^j s_{t+j} = \frac{s_t}{1-\beta}$.[11] We set the discount factor to $\beta = 0.96$. Table A14 shows that using this alternative measure does not materially change our results.

[10] We use Compustat variable `mkvalt` which captures the sum of all issue-level market values, including trading and non-trading issues. We impute missing values with the product of the fiscal-year closing price of the public stock, `prcc_f`, and the number of common shares outstanding, `csho`.

[11] Recall that the time period refers here to an average value within a 5-year window.

Table A15: Relative Bang for the Buck: Median Firm Splits

	Relative Bang & Buck			Bang Components				Buck Components		
Firm group	A_k/A	B_k/B	C_k/C	m_k	$\bar{\epsilon}_k^B$	g_k	θ_k^g	r_k	$\bar{\epsilon}_k^s$	θ_k^s
R&D-int.	0.85	1.37	1.62	0.34	-0.72	0.06	1.51	0.78	-2.46	0.64
Young	1.95	0.90	0.46	0.17	-0.54	0.13	1.39	0.21	-1.68	1.22
Small	2.51	0.25	0.10	0.02	-0.51	0.13	2.75	0.04	-1.89	1.83
- Young	2.74	0.21	0.08	0.02	-0.64	0.16	2.15	0.03	-2.04	1.78
- Old	1.78	0.04	0.02	0.01	-0.23	0.06	7.19	0.01	-1.56	1.94
High-gr.	1.61	1.40	0.87	0.34	-0.85	0.17	0.52	0.42	-1.44	1.05
Random	1.00	1.00	1.00	0.51	-0.36	0.04	2.22	0.49	-1.35	0.99

Note: The table reproduces Table 5 when we split firms by firm shares, rather than sales shares. The market share of "Random" may deviate slightly from 50% due to a finite sample.

D.14 Robustness: Median Firm Splits

In the main text, we classify firms into groups based on their sales shares – requiring $m_k = 0.5$ for all firms. This ensures that all firms have the same potential *capacity* to influence aggregate growth. In this appendix, we redo our analysis using firm shares, rather than sales shares, i.e. each group represents 50% of firms, but not necessarily 50% of sales.

Table A15 shows the results. In this firm-split, small businesses come out as having the largest relative Bang for the Buck. We note, however, that while efficient per dollar spent, the market share of small firms (measuring the capacity to affect aggregate growth) is only 0.02. This highlights that small businesses account for a disproportionately low share of sales.

D.15 Robustness: Common $\psi_i = 2$

In this section we show that our main results are robust to using common $\psi_i = 2$ instead of the one estimated in industry-level regressions. Under this assumption, high-growth firms are slightly more cost-efficient to support than small-young firms. Otherwise, there are no changes to the ranking suggesting that the variation in innovation elasticities is more dominant than that in cost elasticities.

D.16 Robustness: Spillovers with Patent Value-Adjusted R&D Expenses

To quantify technological spillovers between firms, we assume that the extent of spillovers generated by a firm j on any firm i depends on how close to each other these firms are in the technology space (measured by proximity measure $\alpha_{i,j}$ defined in the main text).

Table A16: Decomposition of the Bang for the Buck using common $\psi_i = 2$.

Firm group	Relative Bang & Buck A_k/A	B_k/B	C_k/C	Bang Components m_k	$\bar{\epsilon}_k^B$	g_k	θ_k^g	Buck Components r_k	$\bar{\epsilon}_k^s$	θ_k^s
R&D-int.	0.71	1.01	1.41	0.50	-0.24	0.04	2.72	0.73	-1.17	0.93
Young	1.28	1.43	1.12	0.50	-0.28	0.08	1.75	0.55	-1.27	1.13
Small	1.15	1.27	1.10	0.50	-0.26	0.06	2.24	0.55	-1.21	1.20
- Young	1.36	1.09	0.80	0.34	-0.28	0.08	2.01	0.39	-1.28	1.29
- Old	0.59	0.18	0.30	0.16	-0.04	0.02	18.24	0.16	-0.53	1.86
High-gr.	1.32	1.51	1.14	0.50	-0.59	0.13	0.54	0.58	-1.18	1.02
Random	1.00	1.00	1.00	0.52	-0.27	0.05	1.94	0.51	-1.19	0.92

Note: The table reproduces Table 5 when we let $\psi_i = 2$ for all firms. See also the notes to Table 5. The market share of "Random" may deviate slightly from 50% due to a finite sample.

Table A17: Relative Bang for the Buck including spillovers with patent value as the measure of quality

Firm group	$\frac{A_k}{A}$	$\omega = 0.85$: $\frac{A_k^{spill}}{A^{spill}}$	$\frac{B_k^{int}}{B_k}$	$\frac{B_k^{ext}}{B_k}$	$\omega = 0.75$: $\frac{A_k^{spill}}{A^{spill}}$	$\frac{B_k^{int}}{B_k}$	$\frac{B_k^{ext}}{B_k}$
R&D-int.	0.71	0.69	0.14	0.14	0.68	0.23	0.23
Young	1.38	1.15	0.08	0.04	1.03	0.13	0.06
Small	1.33	1.08	0.05	0.04	0.96	0.09	0.07
High-gr.	1.44	1.23	0.11	0.03	1.12	0.19	0.05
Random	1.00	1.00	0.15	0.16	1.00	0.25	0.26

Note: The table presents the Bang for the Buck and its components when we account for technological spillovers between firms. We use patent values estimated in Kogan et al. (2017) to measure R&D quality. See also notes to Table 6.

Moreover, to accurately gauge the spillovers, we adjust R&D expenses for the quality of innovation a given firm generates. Intuitively, for a given amount of resources devoted to R&D, the more groundbreaking the type of innovation in which a firm is engaged, the more technological spillovers it generates. In the main text, we use citation counts to quantify the quality of R&D expenses in a given firm. In this appendix, we show that the results remain quantitatively unchanged when we use patent value estimated in Kogan et al. (2017) as the measure of R&D quality.

Table A17 reports the results. High-growth firms remain dominant in terms of their relative Bang for the Buck. The ranking of firm groups is the same as in the baseline specification in which we used citation counts to measure R&D quality.

References

Acemoglu, Daron and Dan Cao (2015) "Innovation by Entrants and Incumbents," *Journal of Economic Theory*, 157, 255–294.

Appelt, Silvia, Matěj Bajgar, Chiara Criscuolo, and Fernando Galindo-Rueda (2025) "How effective are R&D tax incentives? Reconciling the micro and macro evidence," CEP Discussion Paper, 2071.

Bloom, Nicholas, Mark Schankerman, and John van Reenen (2013) "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 81 (4), 1347–1393.

de Chaisemartin, Clément and Xavier D'Haultfoeuille (2018) "Fuzzy differences-in-differences," *The Review of Economic Studies*, 85 (2), 999–1028.

Frazzini, Andrea and Lasse Heje Pedersen (2014) "Betting against beta," *Journal of Financial Economics*, 111 (1), 1–25, https://doi.org/10.1016/j.jfineco.2013.10.005.

Gormsen, Niels Joachim and Kilian Huber (2025) "Corporate Discount Rates," *American Economic Review*, 115 (6), 20012049, 10.1257/aer.20231246.

Haltiwanger, John, Ron Jarmin, and Javier Miranda (2013) "Who Creates Jobs? Small Versus Large Versus Young," *Review of Economics and Statistics*, 95 (2), 347–361.

Jaffe, Adam (1986) "Technological Opportunity and Spillovers of R&D: Evidence From Firms' Patents, Profits and Market Value," *American Economic Review*, 76, 984–1001.

Klette, Tor Jakob and Samuel Kortum (2004) "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112 (5), 986–1018.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman (2017) "Technological Innovation, Resource Allocation, and Growth," *The Quarterly Journal of Economics*, 132 (2), 665–712, 10.1093/qje/qjw040.

Lentz, Rasmus and Dale T Mortensen (2008) "An empirical model of growth through product innovation," *Econometrica*, 76 (6), 1317–1373, Publisher: Wiley Online Library.

Miles, James A. and John R. Ezzell (1980) "The Weighted Average Cost of Capital, Perfect Capital Markets, and Project Life: A Clarification," *The Journal of Financial and Quantitative Analysis*, 15 (3), 719–730, `http://www.jstor.org/stable/2330405`.

Modigliani, Franco and Merton H. Miller (1963) "Corporate Income Taxes and the Cost of Capital: A Correction," *The American Economic Review*, 53 (3), 433–443, `http://www.jstor.org/stable/1809167`.

Mukoyama, Toshihiko and Sophie Osotimehin (2019) "Barriers to Reallocation and Economic Growth: The Effects of Firing Costs," *American Economic Journal: Macroeconomics*, 11 (4), 235–270.

Ottonello, Pablo and Thomas Winberry (2024) "Capital, Ideas, and the Costs of Financial Frictions," mimeo.

Thomson, Russell and Elizabeth Webster (2012) "The Design of R & D Support Schemes for Industry," *Economic Papers: A journal of applied economics and policy*, 31 (4), 464–477.