Abstract

This paper analyzes the implications of advertising for firm dynamics and economic growth in the long run through its interaction with R&D investment at the firm level. We develop a model of endogenous growth with firm heterogeneity that incorporates advertising decisions. We model advertising by constructing a framework that unifies a number of facts identified by the empirical marketing literature. We then calibrate the model to match several empirical regularities across firm sizes using U.S. data. Through a novel interaction between R&D and advertising, we are able to explain the empirically observed deviation from size-invariant firm growth rates (Gibrat’s law) as well as the behavior of R&D expenditures across firm size. In addition, our model predicts a substitution effect between R&D and advertising at the firm level. Lower advertising costs are associated with lower R&D investment, slower growth and lower welfare. We provide empirical evidence supporting the substitution between R&D and advertising investment using exogenous variation in the cost of R&D arising from changes in the tax treatment of R&D expenditures over time and across U.S. states. We study the policy implications of our model in terms of advertising tax and R&D subsidies. Taxing advertising is shown to have a positive but relatively small effect on economic growth within a reasonable range of tax rates. We find that R&D subsidies are more effective under an economy calibrated to include advertising relative to one with no advertising.

JEL codes: E2; L1; M3; O31, O32, O33, and O41.

Keywords: Endogenous growth; Advertising; Innovation; Research and Development; Firm Dynamics; Policy.

*We are indebted to Jess Benhabib for his invaluable advice and support, and to Gian Luca Clementi, Boyan Jovanovic, and Edouard Schaal for continued discussions and suggestions. For helpful comments, we thank Diego Anzoategui, Salomé Baslandze, Alberto Bisin, Jaroslav Borovička, Katarína Borovičková, Diego Daruich, Tülin Erdem, Jordi Gali, Gino García, Sonia Gilbukh, Seher Gupta, Siddharth Hari, Deepak Hegde, Masakazu Ishihara, Olga Itenberg, Nicholas Kozeniauskas, Julian Kozlowski, Pedro Mendi, Simon Mongey, Ilari Paasivirta, Diego Perez, Michael Peters, Xavier Ragot, Felipe Saffie, Tom Schmitz, Venky Venkateswaran, Jaume Ventura and Gianluca Violante. We also received valuable feedback from seminar participants at New York University, and from conference participants at the 2016 North American Summer Meeting of the Econometric Society (Philadelphia), EEA-ESEM 2016 (Geneva), SED 2016 (Toulouse), and the XXXI Jornadas de Economía Industrial (Palma de Mallorca). Any remaining errors are our own.

†Job Market Candidate. Department of Economics, New York University, 19 W 4th street, 6th floor, New York, NY 10012. Email: laurent.cavenaile@nyu.edu

‡Department of Economics, New York University. Email: pau.roldan@nyu.edu
1 Introduction

Since the seminal contributions of Romer (1990), Grossman and Helpman (1991) and Aghion and Howitt (1992), the economic literature has widely emphasized the role of innovation in the process of economic growth. Research and Development (R&D) introduces goods of higher quality and enhanced production technologies which raise living standards. From a firm’s perspective, innovation is used strategically to increase profits. By performing R&D, firms can increase their market share by selling higher quality goods and diverting demand from lower-quality products. While this process and its implications for economic growth are relatively well acknowledged in the literature, there exist other tools that firms use to shift demand towards their products. One obvious and non-negligible example among firm intangibles is advertising.¹

By advertising their goods, firms can alter consumer preferences and ultimately increase demand for their products. This suggests that innovation and advertising may be substitutable tools in the quest of firms for higher profits. On the other hand, advertising investment could also raise the return to innovation by increasing market shares of new products. This implies that advertising decisions are potentially not neutral in terms of global economic growth. In fact, both these tools represent sizeable expenditures in the aggregate economy. Over the 1980 – 2013 period in the U.S., the share of R&D expenditures over GDP fluctuated between 2.27% and 2.82%.² Over the same time period, firms in the U.S. spent on average around 2.2% of GDP on advertising each year.³ Yet, the growth literature is relatively silent on the potential interaction between R&D and advertising expenditures and its potential impact on firm growth, firm dynamics and long run economic development.

In this paper, we ask how advertising affects R&D investment decision at the firm level and study the implications for welfare, growth and firm dynamics. We show that a decrease in the relative cost of advertising can have significant detrimental effects on long run economic growth through a substitution

¹A recent trend in the macroeconomic literature has started investigating several types of intangible investment and their potential implications for the overall economy. For instance, McGrattan (2015), McGrattan and Prescott (2014) and Gourio and Rudanko (2014a) study how intangible investment may affect productivity measurement and business cycle fluctuations, while Hall (2013) and Molinari and Turino (2009) specifically focus on advertising. Gourio and Rudanko (2014b) and Arkolakis (2010, 2016) consider the impact of marketing on firms’ sales, investment and exports in models with frictional product markets in which firms must use marketing to reach consumers. Atkeson and Kehoe (2005) focus on organization capital as a form of intangible investment.


³Source: Coen Structured Advertising Expenditure Dataset, extracted from the McCann Erikson advertising agency, running until 2007 (available at http://www.purplemotes.net/2008/09/14/us-advertising-expenditure-data/). These data include advertising expenditures on TV, newspapers, radio, magazines, telephone directories, the Internet, direct mail, billboard and outdoor advertising, and business papers (trade press). The data do not include other forms of marketing, for instance brand sponsorship, sales promotion, or interactive marketing.
effect between R&D and advertising at the firm level. We also find that the interaction between R&D and advertising can generate some observed empirical regularities related to firm dynamics, such as decreasing firm growth rates with firm size (a deviation from the so-called Gibrat’s law) as well as the decreasing R&D intensity with firm size.

We build on the Akcigit and Kerr (2015) model of endogenous growth through R&D which we extend to incorporate explicit advertising decisions. Firms are heterogenous and own different portfolios of goods that they monopolistically supply. Product quality grows on a ladder through innovation arising from investment in R&D, which can take two forms in our model. Through internal R&D, firms can increase the quality of their own goods. External R&D enables incumbent firms and potential entrants to improve on the quality of a good that they do not own and displace the former producer through creative destruction. Besides R&D, firms can use advertising to expand their market shares and profit. While the intrinsic quality of goods in the economy moves along a quality ladder only through innovation, firms use advertising to influence the perception that consumers have of these goods. In our model, advertising alters consumer preferences and ultimately acts as a demand shifter. Even though both R&D and advertising in our model can both affect market shares, long term growth only comes from R&D and innovation.

We calibrate the model via indirect inference using both aggregate and micro-level data. In particular, our model is calibrated to match four empirical facts related to firm dynamics and investment across firm size. Figure 1 shows these four facts. The upper-left panel displays the average growth rate of publicly-listed firms in the U.S., by size quintiles. Growth rates are higher for smaller firms and are consistently decreasing with firm size. The upper-right panel shows that small firms also tend to be relatively more R&D intensive. This suggests that smaller firms experience higher growth through higher R&D investments, a conclusion that has sometimes been interpreted in the existing literature as evidence of technological differences in R&D production across firm size. This interpretation has important implications in terms of policy recommendations (e.g. Akcigit (2009), Acemoglu, Akcigit, Bloom, and Kerr (2013) or Akcigit and Kerr (2015)). If small firms are more efficient at innovating, R&D subsidies targeted at small firms may be optimal. An important question is then to understand the source of the deviation from constant R&D intensity across firm size. In this context, we show

4Quintiles are based on sales normalized by total sales in the same year. Data include U.S. listed companies performing R&D and advertising between 1980 and 2015 from the Compustat database. More details on the sample selection in Section 3.

5R&D intensity is measured by R&D expenditures as a percentage of sales.
that the observation that R&D intensity diminishes with firm size can arise as the result of the optimal allocation of resources between R&D and advertising at the firm level.

Our calibration is also disciplined by two new facts about advertising. First, we find a negative relationship between advertising intensity (i.e., advertising expenditures normalized by size) and firm size, which can be observed in the lower-left panel of Figure 1. Second, we find that larger firms rely relatively more on advertising compared to R&D, as the ratio of R&D to advertising expenditures decreases with firm size (lower-right panel in Figure 1). This suggests that firms choose to change the mix of R&D and advertising and switch to relatively more advertising as they grow larger.

Our calibration delivers two main implications regarding the interaction between R&D and advertising. First, our results show the existence of a substitution effect between R&D and advertising at the firm level. A decrease in the cost of advertising leads to an increase in the entry rate and in creative destruction. This, in turn, decreases the incentive for incumbent firms to invest in R&D and shifts the firm size distribution to the left (more small firms). Overall, investment in R&D decreases which leads to slower economic growth and lower welfare. If advertising is banned, our model predicts an increase in the growth rate of the economy by 0.64 percentage point. Second, regarding firm dynamics, we find that the interaction between R&D and advertising can quantitatively deliver the four facts reported in

![Figure 1: Firm growth, R&D intensity, advertising intensity, and R&D-advertising ratio, across firm size quintiles.](image)

Notes: Firms are ranked in size quintiles according to their normalized level of sales (sales as a ratio of average sales in the same year). R&D and advertising intensities are measured as the ratio of total R&D and advertising expenditures to total sales within each group.
Figure 1, which suggests that the deviation from constant R&D intensity across firm size should not necessarily be interpreted as evidence for differences in terms of R&D efficiency.

The main new mechanisms in our model are based on empirical observations from the marketing literature. There is ample evidence in the empirical marketing literature that larger firms have a cost advantage in terms of advertising compared to firms with fewer products. One reason for this size advantage comes from a spillover effect through different goods under the same brand name. By advertising one product under a given brand name, a firm can influence not only the perception of the quality of the advertised good but also of other goods sharing the same brand name. The existence of this spillover effect affects firms’ dynamic incentives to engage in R&D. As smaller firms gain relatively more in terms of advertising spillover from acquiring an additional product, they optimally choose to perform relatively more R&D in order to expand to new product markets. As a result, these firms grow relatively faster than large ones.

Our calibration also matches several untargeted moments among which the decrease in the variance of firm growth as well as of R&D and advertising intensity with firm size. Moreover, using exogenous variation in the cost of R&D arising from changes in the tax treatment of R&D expenditures across U.S. states and over time, we provide empirical evidence supporting the substitution between R&D and advertising which is at the core of the predictions of our calibrated model.

We study policy implications of our model. First, we show that taxes on advertising, while desirable from the angle of growth, might have a relatively small impact. Second, we show that identifying the source of the observed decrease in R&D intensity with firm size is relevant for R&D policy. R&D subsidies are more effective at promoting economic growth if the deviation from constant R&D intensity comes from the interaction between R&D and advertising rather than from technological differences in terms of R&D efficiency across firm size (the increase in growth is 0.13 percentage point higher in the former case for a 50% subsidy).

The rest of the paper is organized as follows. Section 2 presents a comprehensive review of the literature with a special focus on endogenous growth, firm dynamics and the marketing literature used to motivate our advertising modeling approach. Section 3 presents key empirical facts related to Gibrat’s law, R&D and advertising intensity at the firm level, as well as the relative use of R&D and advertising across firm size. We then describe the model and discuss its comparative statics in Section 4. Section 5 presents the calibration and shows that our novel advertising mechanism is able to quantitatively replicate the main empirical regularities described in Section 3. We also
utilize the calibrated model to ask how advertising affects the firm size distribution, firm entry rates, economic growth and welfare. In Section 6 we present a set of validation exercises, including the model performance in terms of untargeted moments (in Section 6.1), and empirical evidence supporting the main mechanism at work behind the calibrated model (in Section 6.2). Section 7 investigates policy implications regarding advertising taxes and R&D subsidies. Section 8 concludes. The Appendix includes proofs of all the results, model extensions mentioned in the main text, and additional tables and figures.

2 Related Literature

Our work is related to several strands of the literature. First and foremost, we build upon models of endogenous firm growth through product innovation. This area was pioneered by Klette and Kortum (2004), who built a stylized version of the Schumpeterian creative-destruction models of Grossman and Helpman (1991) and Aghion and Howitt (1992) into a model of multi-product firm dynamics. Their work is able to exhibit several behaviors that are consistent with empirical finding coming from micro-level data, especially regarding the right-skewness in the firm size distribution, the persistence in firms’ R&D investments, and the volatility of innovation in the cross-section of firms.\(^6\)

However, by assuming that firm productivity scales perfectly with size, this framework fails to explain why firms of different sizes may grow at different rates. In fact, their model delivers a theoretical version of Gibrat’s law (from Gibrat (1931)), in that firm growth is independent of size. Recent empirical studies have argued that there exist deviations from Gibrat’s law especially among continuing establishments. Conditional on survival, small establishments grow faster than large ones, and net exit rates are a decreasing function of size.\(^7\)

One contributor to such phenomenon is that smaller firms tend to be more innovation-intensive.\(^8\)

Accordingly, a wave of second-generation models of innovation-driven firm growth has emerged from

---

\(^6\)The applied work by Lentz and Mortensen (2008) further unveiled the predictive power of the Klette and Kortum (2004) model by putting it to a rigorous empirical test.

\(^7\)See Hall (1987), Sutton (1997), Caves (1998), and Geroski (1998), all of whom argue that growth rates of surviving firms tend to systematically decrease with increasing firm size. Audretsch, Klomp, Santarelli, and Thurik (2004) argue that deviations from Gibrat’s law are especially prevalent within the manufacturing industry, and Rossi-Hansberg and Wright (2007) find similarly that U.S. sectors with larger physical capital shares exhibit significantly stronger scale dependence in establishment dynamics and distributions. For a survey of this empirical literature, see Santarelli, Klomp, and Thurnik (2006).

\(^8\)Cohen and Klepper (1996) show that small innovating firms generate more innovations per dollar spent in R&D, and Akcigit and Kerr (2015) show that small firms spend more in R&D per dollar of sales. Using Compustat data, we confirm some of these results in Section 3.
subsequent applied work by Acemoglu, Akcigit, Bloom, and Kerr (2013), Acemoglu and Cao (2015) and Akcigit and Kerr (2015), among others. These papers have proposed extending the Klette and Kortum (2004) framework to include heterogeneity in innovation technologies in order to incorporate a more meaningful interaction between different types of research (e.g., product versus process, radical versus incremental), as well as between entrants and incumbent firms. Most of the theoretical approaches to date impose parametric restrictions on the innovation sector in order to limit growth and feature size-dependence in equilibrium.\footnote{For instance, Akcigit and Kerr (2015) assume weak scalability in the innovation technology that is used most intensively by small firms; Acemoglu and Cao (2015) assume that innovations by small firms lead to higher step sizes of technical advance; and Acemoglu, Akcigit, Bloom, and Kerr (2013), who assume an exogenous stochastic process for firms’ productive capacity, impose that lower capacity states are highly absorbing so that most large surviving firms belong to the low state.}

In contrast, we propose a theory that, while building upon the aforementioned work, does not hinge on specific technological size-dependence in R&D. Yet, our theory is able to deliver that smaller firms spend in equilibrium a higher fraction of sales to R&D expenditures, and that those firms grow faster as a result.

Our novel trade-off between innovation and advertising decisions of firms introduces the required non-homogeneities in equilibrium. Specifically, we rely on two main strands of the marketing literature in order to build a motive for advertising into our endogenous growth framework.

First, in our model, larger firms experience higher returns to advertising expenditures. A long tradition in the marketing literature identifies this phenomenon as the equity value of so-called “umbrella-branding”. It suggests that there exist spillovers between goods within the firm, in the sense that increasing advertising expenditures on one good not only increases sales for that good but also indirectly for all other goods under the same brand.\footnote{In the model, we take as given that each firm has a unique brand, and we abstract from brand choice.} Lane and Jacobson (1995) and Tauber (1981, 1988) show that brand developments can decrease marketing costs, and Rangaswamy, Burke, and Oliver (1993) show that they can enhance marketing productivity. Moreover, Smith and Park (1992) show that they can help capture greater market share. Using household scanner panel data for the U.S., Balachander and Ghose (2003) find a positive and significant spillover effect for multiple product categories and geographic markets. Dacin and Smith (1994) show, by means of controlled experiments, a positive relationship between the number of products affiliated with a brand and consumers’ confidence in the quality extension of the brand. Erdem (1998) argues empirically that the quality perceptions of a brand in a product category are positively affected by the consumer’s experience with the same brand in a different category, because branding allows consumers to learn faster about quality through use experience. More particularly for our purposes, Erdem and Sun (2002) show that the effects identi-
fied in Erdem (1998) further translate to positive spillovers between advertising and sales for different goods within the same umbrella brand. Specifically, such effects are also present in innovation-driven industries, e.g. the pharmaceutical industry (see Suppliet (2015)). More generally, Ailawadi, Harlam, César, and Trounce (2006), Leeflang, Parreno-Selva, Dijk, and Wittink (2008), Bezawada, Balachander, Kannan, and Shankar (2009), and Leeflang and Parreno-Selva (2012) all show significant cross-category sales and advertising dependencies for a variety of retail markets.

The second observation that we draw from the marketing literature is that higher-quality goods within the firm benefit relatively more from advertising: everything else equal, a good that compares better to other goods within the brand in terms of quality experiences higher increases in sales per dollar spent into advertising. Archibald, Haulman, and Moody (1983) and Caves and Greene (1996) find a positive correlation between advertising and quality for innovative goods as well as for experience goods, i.e. goods for which the experience of buyers is useful in making brand choices.11 Tellis and Fornell (1988) examine the relationship over the product life cycle, and find that advertising and profitability are both positively correlated to quality. Further, Homer (1995), Kirmani and Wright (1989) and Kirmani (1990, 1997) offer experimental evidence suggesting that the positive relationship exists because consumers perceive higher advertising expenditures indicating that the good is of high quality.12

More broadly, our theoretical approach is also consistent with standard modeling in both Marketing and Economics. On the one hand, the marketing literature typically argues that advertising firms value “goodwill”, namely the stock of advertising accumulated over time. The most extended approach consists of posing a logit discrete-choice model of demand, coupled with an explicit law of motion for goodwill.13 In the empirical application, most studies find a large depreciation rate for goodwill implying that, over a one-year period, goodwill almost fully depreciates. In our model, we assume that advertising has an instantaneous effect on firms’ operating profits, while all dynamic effects operate through innovation. In Appendix C.2.1 we show an extension of the model that accommodates explicit goodwill accumulation within our baseline model. We further show that allowing for goodwill accumulation would not qualitatively change the results of our baseline model.

Regarding advertising in Economics, our paper relates to a long tradition of studying advertising

---

11In this empirical literature, quality is usually proxied by consumer reports and rankings of consumer satisfaction.
13See Dubé, Hitsch, and Manchanda (2005) and Doganoglu and Klapper (2006). For direct applications of this approach in Macroeconomics, see Nervole and Arrow (1962) and Molinari and Turino (2009).
as an explicit factor affecting consumer tastes, as in Dorfman and Steiner (1954), Dixit and Norman (1978), Becker and Murphy (1993) and Benhabib and Bisin (2002). Our approach is similar to those in that firms’ advertising effectively acts as a demand shifter in equilibrium. A parallel body of literature introduces a role for valuable customer capital into macro models, viewing marketing as a tool to build continuing buyer-seller relationships for firms due to either the existence of frictions in product markets (e.g. Gourio and Rudanko (2014b), Hall (2008)) or because of costs to market penetration (e.g. Arkolakis (2010, 2016) and Eaton, Eslava, Jinkins, Krizan, and Tybout (2014)). Perla (2015) views advertising as a way of raising product awareness among customers, but his paper focuses on the implications of product sorting by consumers on the industry life cycle and the degree of market competition, while we focus on the interaction between advertising and innovation and its effects on growth. Fishman and Rob (2003) study the implications of customer accumulation for firm size and R&D choices, but they also do not focus on how marketing and innovation interact. To our knowledge, only Grossmann (2008) has considered this potential interaction from a purely theoretical perspective. However, while in that paper advertising and R&D are assumed to be direct complements, we opt to be agnostic about the degree of complementarity between the two, and instead discipline their substitution by matching directly firm-level cross-sectional data. In addition, we incorporate firm heterogeneity which allows us to study the implications of advertising for firm dynamics. Moreover, our policy analysis puts a special emphasis on how advertising alters both the R&D-vs-advertising composition of firm investment on the intensive margin, and the size distribution of firms on the extensive margin. In the calibrated economy, we find that advertising is detrimental to long run growth and welfare even though it may contemporaneously expand demand.

3 Empirical Findings

In this section, we present empirical results on firm growth, R&D and advertising expenditures. Some of the coefficients that we obtain will then be used to calibrate the model that is presented in Section 4.

Data We use annual data on U.S. listed companies from Compustat over the period 1980-2015. Firms reporting nonpositive sales or nonpositive employment are excluded from the sample. We also exclude non-innovative (i.e., no R&D expenditures) and non-advertising (i.e., no advertising expenditures) firms. To exclude outliers, we ignore firms experiencing year-on-year sales growth rates of more than
1,000%, as well as R&D-to-sales and advertising-to-sales ratios of more than 100%.\textsuperscript{14} Moreover, we exclude mergers and acquisitions, and in order to correct for a possible survival bias in the growth regressions, we make the conservative assumption that exiting firms have a growth rate of \(-100\%\) in their last period of operation. Basic descriptive statistics of the resulting sample are reported in Table A.1 of Appendix A.

**Empirical Facts and Relation to the Literature**  
Our main focus is on the regression results for the four empirical facts that we presented in Figure 1.\textsuperscript{15} We use firm sales as our baseline measure of firm size.\textsuperscript{16} All four regressions are of the form:

\[
y_{ij,t} = \alpha_0 + \beta_1 \log(Sales_{ij,t}) + \beta_2 FirmAge_{ij,t} + \beta_3 FinConst_{ij,t} + \alpha_j + \alpha_t + u_{ij,t}
\]

for firm \(i\) in industry \(j\) in year \(t\), where \(\alpha_t\) controls for time fixed effects and \(\alpha_j\) controls for industry fixed effects and where the dependent variable is

\[
y_{ij,t} \in \left\{ \frac{\Delta Sales_{ij,t}}{Sales_{ij,t}}, \log \left( \frac{R&D_{ij,t}}{Sales_{ij,t}} \right), \log \left( \frac{Adv_{ij,t}}{Sales_{ij,t}} \right), \log \left( \frac{R&D_{ij,t}}{Adv_{ij,t}} \right) \right\}
\]

for each of the four facts, respectively. To control for firm characteristics, we include firm age, which we measure as the number of years elapsed since the firm first appeared in the sample,\textsuperscript{17} and a measure of financial constraints, which is the difference between sales and purchases of common and preferred stock as a share of firm sales.\textsuperscript{18}

In the results that follow, Facts \#1 and \#2 (related to firm growth and R&D intensity, respectively) have already been highlighted in the literature. In addition, we also obtain Facts \#3 and \#4 (related to advertising intensity and the R&D-advertising ratio, respectively), which are new to our work.

**Fact \#1:** *Smaller innovative firms grow faster on average.*

Table 1 shows that there exists a significant deviation from Gibrat’s law among innovative firms. The results show that larger innovative firms tend to experience lower growth rates in terms of sales.

---

\textsuperscript{14}Using different thresholds does not change the sign nor the significance of the coefficients of interest.

\textsuperscript{15}Regression results in Table A.2 of Appendix A confirm that both R&D and advertising increase sales and profit.

\textsuperscript{16}Table A.3 of Appendix A shows results for all four facts when size is measured by assets and employment. We find that the results do not change qualitatively.

\textsuperscript{17}This measure of firm age in the Compustat database is standard and has been used among others by Shumway (2001), Lubos and Veronesi (2003), Fama and French (2004) and Chun, Kim, Morck, and Yeung (2008).

\textsuperscript{18}We borrow this measure from Itenberg (2015).
In our full specification (column (3)), a 1% increase in sales translates into a 0.0325% decrease in sale growth.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(Sales/Sales)</td>
<td>-0.0436***</td>
<td>-0.0338***</td>
<td>-0.0325***</td>
</tr>
<tr>
<td></td>
<td>(0.00275)</td>
<td>(0.00282)</td>
<td>(0.00288)</td>
</tr>
<tr>
<td>Firm Age</td>
<td>-0.00474***</td>
<td>-0.00441***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000368)</td>
<td>(0.000367)</td>
<td></td>
</tr>
<tr>
<td>Fin. Const.</td>
<td>0.00270</td>
<td></td>
<td>0.00270</td>
</tr>
<tr>
<td></td>
<td>(0.00172)</td>
<td></td>
<td>(0.00172)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.346***</td>
<td>0.322***</td>
<td>0.313***</td>
</tr>
<tr>
<td></td>
<td>(0.0438)</td>
<td>(0.0409)</td>
<td>(0.0454)</td>
</tr>
<tr>
<td>Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Industry FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>27191</td>
<td>27191</td>
<td>24856</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 1: Firm growth regressions (Fact #1).
Notes: Compustat data for 1980-2015. The sample is restricted to firms reporting strictly positive sales, strictly positive employment, strictly positive R&D and advertising expenditures, with year-on-year sales growth less than 1,000%, and R&D-to-sales and advertising-to-sales ratios of less than 100%. Mergers and acquisitions have been excluded from the sample. Age is measured as the elapsed time since the first observation in the data. Our measure of financial constraints is sale minus purchases of common and preferred stock, divided by sales. Sales and advertising expenditures are in thousands of U.S.D. Standard errors are clustered by firm (in parentheses). Significance level: * 10%; ** 5% ; *** 1%.

These results provide evidence that is similar to that reported in Akcigit (2009) and Akcigit and Kerr (2015).\textsuperscript{19} The phenomenon has also been studied in the literature within different theoretical frameworks. Models of firm dynamics following Hopenhayn (1992), based on stochastic (Markov) productivity shocks, can replicate the negative correlation between firm size and firm growth. If the set of productivity shocks is finite and mean-reverting, larger firms (i.e., firms with higher productivity) eventually face a limit to growth and experience lower average growth rates than small firms. Jovanovic (1982) proposes a model of learning in which firms receive noisy signals about their productivity (or cost function) and update their beliefs accordingly. Firms with beliefs below a certain threshold optimally choose to exit. Younger firms are more uncertain about their productivity and hence learn more than older firms. This leads to a stronger revision of their beliefs and hence higher growth rates (among surviving firms).\textsuperscript{20} On average, younger firms grow faster and are larger than older ones. Thus, the

\textsuperscript{19}A large empirical literature provides evidence of a deviation from Gibrat’s law (see Santarelli, Klomp, and Thurnik (2006) for a review of the literature).

\textsuperscript{20}In the long run, surviving firms learn their true productivity and hence do not grow any more.
correlation between firm size and firm growth comes through an age effect and the positive correlation between age and size. Consequently, in this type of models, the observed association between firm size and firm growth should vanish as one controls for firm age. From an empirical perspective, Haltiwanger, Jarmin, and Miranda (2013) argue, using data about the whole universe of firms in the U.S., that the observed deviation from Gibrat’s law does not systematically hold when one controls for firm age. In column (2), we find that the significant deviation from Gibrat’s law still holds after controlling for firm age in our sample of innovative firms.

In another strand of literature, Cooley and Quadrini (2001) link the dependence of firm growth on size and age to financial market frictions. In a model of firm dynamics similar to that of Hopenhayn (1992), Cooley and Quadrini (2001) show that the interaction of persistent productivity shocks with costly equity issuance and default generates a higher financial leverage for small and young firms and qualitatively replicates the decreasing relationship between firm size, firm age and expected growth. In column (3) of Table 1, we find that the deviation from constant growth holds after controlling for financial constraints.

**Fact #2:** Smaller innovative firms have higher R&D intensity on average.

An emerging trend in the economic growth literature has investigating the departure from constant growth across firm size by linking it to another cross-sectional empirical fact: the higher R&D intensity of smaller firms (see, for example, Akcigit (2009), Acemoglu and Cao (2015), and Akcigit and Kerr (2015)). As larger firms invest relatively less in innovation, they experience relatively lower growth rates. Acemoglu, Akcigit, Bloom, and Kerr (2013) and Acemoglu, Akcigit, and Celik (2014) also propose models in which older and larger firms invest relatively less in R&D and grow more slowly as a result.

Table 2 shows that this pattern holds in our sample as well: larger firms are less R&D intensive (R&D intensity is measured by R&D expenditure as a share of sales). In addition, we control for firm age and our measure of financial constraints and find that the negative relationship between R&D

---

21Second-generation models building on these theories have since emerged attempting to break this conditional independence. For instance, Clementi and Palazzo (2016) propose an extension of the Hopenhayn (1992) model with capital accumulation in which, conditional on size, growth rates are on average decreasing in age because the presence of adjustment costs to capital implies cross-sectional dispersion in productivity for given levels of capital, with highly-productive growing firms being younger on average.

22See also Clementi and Hopenhayn (2006). Itenberg (2015) links the high innovation intensity of small firms to the extensive use of external equity financing of these firms. For a related approach, see Schmitz (2015). Similarly, Garcia-Macia (2015) shows using Spanish data that young firms tend to be leveraged and are intensive in intangible assets, whereas older larger firms tend to deleverage as they age further and face less stringent financial constraints.
intensity and firm size still holds. The effect of age on R&D intensity is positive though not strongly significant, and it does not alter the sign or significance of the size effect. In our full specification (column (3)), a 1% increase in sales translates into a 0.1035% decrease in the R&D expenditure to sales ratio.

In sum, smaller innovative firms grow relatively faster, and these firms spend relatively more in R&D. While we are not the first ones to make these observations using Compustat data (see, for instance, Akcigit (2009) and Itenberg (2015)), we show in addition that size remains significant after controlling for age and financial constraint both in growth and R&D intensity regressions, which suggests that explanations based on the firm’s product life-cycle and financial constraints cannot entirely account for the deviation from Gibrat’s law observed in the data. We propose an alternative explanation for this deviation based on the interaction between R&D and advertising in firms’ optimal decisions. For this, we present two new facts which are used to calibrate our model.

**Fact #3:** *Smaller innovative firms have higher advertising intensity on average.*

Table 3 shows that advertising intensity is decreasing in firm size: smaller firms spend more in advertising per dollar of sales. Once again, this correlation remains significant after controlling for age and financial constraints. In our full specification, we find that a 1% increase in sales translates into a 0.0317% decrease in the advertising expenditure to sales ratio.
Table 3: Advertising intensity regressions (Fact #3). Notes: See Table 1.

Fact #4: Smaller innovative firms have higher R&D to advertising ratios on average.

Finally, in Table 4, we find that larger firms tend to use relatively more advertising than R&D, as the ratio of R&D to advertising expenditure is significantly decreasing in size. In particular, a 1% increase in sales leads to a 0.0719% decrease in the R&D to advertising expenditure ratio. This suggests that as innovative firms grow larger, they tend to substitute R&D for advertising. We also find that the effect of age on the relative use of R&D and advertising is not robustly significant.

Overall, these results suggest that the interaction between R&D and advertising can potentially affect firm growth through a size effect, consistent with the advertising spillover that our model captures. In addition, our estimated model will deliver all four of these predictions in equilibrium.

4 An Endogenous Growth Model with Advertising

Inspired by the empirical findings of the previous section, we build a theory of advertising into a model of endogenous growth. The basic structure follows the framework set up by Klette and Kortum (2004), and specifically allows for heterogeneous innovations as in the recent work by Akcigit and Kerr (2015). The advertising theory, on the other hand, is based on empirical regularities that we take from the marketing literature. Our broader goal is to illustrate a novel interplay between innovation and
Table 4: R&D to advertising expenditure regressions (Fact #4). Notes: See Table 1.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(Sales)</td>
<td>-0.0657***</td>
<td>-0.0703***</td>
<td>-0.0719***</td>
</tr>
<tr>
<td></td>
<td>(0.0105)</td>
<td>(0.0116)</td>
<td>(0.0120)</td>
</tr>
<tr>
<td>Firm Age</td>
<td>0.00223</td>
<td>0.00227</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00214)</td>
<td>(0.00214)</td>
<td></td>
</tr>
<tr>
<td>Fin. Const.</td>
<td></td>
<td></td>
<td>-0.00207</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.00498)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.455***</td>
<td>1.467***</td>
<td>1.432***</td>
</tr>
<tr>
<td></td>
<td>(0.223)</td>
<td>(0.225)</td>
<td>(0.236)</td>
</tr>
<tr>
<td>Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Industry FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>27191</td>
<td>27191</td>
<td>24856</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.44</td>
<td>0.44</td>
<td>0.45</td>
</tr>
</tbody>
</table>

advertising, which critically shapes firms’ dynamic incentives for growth and, thereby, growth in the economy as a whole.

4.1 Environment

4.1.1 Preferences

Time is continuous, infinite, and indexed by $t \in \mathbb{R}_+$. The economy is populated by a measure-one of identical, infinitely-lived individuals with discount rate $\rho > 0$. A representative household has preferences given by:

$$U = \int_0^{+\infty} e^{-\rho t} \ln (C_t) dt \tag{1}$$

where $C_t$ is consumption of the single final good, whose price is normalized to $P_t = 1, \forall t$. The household is endowed with one unit of time every instant, which is supplied inelastically to the productive sector of the economy in the form of labor. The wage rate is denoted by $w_t$ and is determined endogenously to clear the labor market. The household owns all the firms in the economy and carries each period a stock of wealth $A_t$, equal to the total value of corporate assets. Wealth earns an
instantaneous and time-varying rate of return $r_t$. The flow budget constraint is, therefore:

$$\dot{A}_t = r_t A_t + w_t - C_t$$

(2)

where $A_0 \geq 0$ is given.

**4.1.2 Final Good Sector**

The final good is produced by a representative final good firm using a measure-one continuum of input varieties, indexed by $j \in [0, 1]$, which are supplied by an intermediate good sector. Technology is given by the Cobb-Douglas production function:

$$Y_t = \frac{1}{1-\beta} \int_0^1 q^{\beta} y_{jt}^{-\beta} dj$$

where $\beta \in (0, 1)$. Here, $y_{jt}$ is the quantity of intermediate good $j$ that is used at time $t$. Quantities are weighted by the term $q_{jt}$, which stands for the quality of the good that is being perceived by agents in the economy.

**4.1.3 Product Qualities**

The total perceived quality $q_{jt}$ of a certain good $j$ at time $t$ is defined by:

$$q_{jt} = q_{jt}(1 + d_{jt})$$

Total perceived quality includes two components. The first component is referred to as the intrinsic quality of the product, denoted by $q_{jt} > 0$, and it stands for the currently leading-edge technological efficiency of the product. This component of quality is built upon over time through a process of technical innovation. In particular, it advances on a ladder, as in the models of Grossman and Helpman (1991) and Aghion and Howitt (1992). This phenomenon is induced by expenditure into R&D at the productive sector level in a way that will be described shortly.

The second component in total perceived quality is the so-called extrinsic quality of the product,

23 A more general specification for preferences is $Y_t = \frac{1}{1-\beta} X_t^{1-\beta}$, where $X_t = \left( \int_0^1 q^{\beta} y_{jt}^{-\beta} dj \right)^{1-\beta}$ is a CES aggregator of input varieties and $\epsilon > 1$ is the elasticity of substitution between input varieties. Assuming $\beta = \epsilon^{-1}$ gives us the Cobb-Douglas aggregator assumed in the main text. Appendix C.2.6 shows that there is no loss in terms of the qualitative properties of the model by making this simplifying assumption.
given by $\phi_{jt} \equiv q_{jt}d_{jt}$, where $d_{jt} \geq 0$ denotes advertising-induced quality. This component of total quality refers to the part of total perceived quality of a good that is induced by the producer’s advertising efforts on that specific product. Therefore, we may also refer to $\phi_{jt}$ as the effectiveness of advertising product $j$ in period $t$. Besides advertising expenditures, $\phi_{jt}$ is also a function of the size of the firm and advertising efficiency, in a way that we specify in detail in the next section.

A few comments on our specification of advertising are in order. First, note that intrinsic ($q_j$) and extrinsic ($\phi_j$) quality are substitutes at the good level, as they enter additively into total quality (i.e., $\bar{q}_j = q_j + \phi_j$). While spending on advertising (which increases $\phi_{jt}$) is not indispensable for a good to be perceived with some positive quality by consumers, it does create an added value over and above the quality level that is purely intrinsic to the good. However, we also allow for a degree of complementarity coming from the fact that $\phi_{jt}$ is itself an increasing function of intrinsic quality, $q_j$. Namely, advertising is effective in raising perceived quality only if the good has nonzero intrinsic quality, and it is more effective the higher the intrinsic quality of the good.

Second, we model advertising as a preference shifter. By choosing their advertising expenditures, firms are effectively impacting perceived qualities and, therefore, consumer preferences. Indeed, as will become apparent later, the endogenous choice of advertising expenditures by firms serves as a consumer demand-curve shifter in equilibrium. Our specification follows Dixit and Norman (1978), Becker and Murphy (1993), Benhabib and Bisin (2002) and Molinari and Turino (2009), among others, who model advertising through product-specific taste parameters or as an explicit argument in the utility function. In Appendix C.2, we present various alternative ways to model advertising and show that, in all cases, advertising appears as a demand shifter.

Third, we adopt the view that advertising is persuasive, and not purely informative. By this we mean that consumers cannot choose what information to be exposed to, but rather behave according to the shifts in tastes induced by advertising, which they take as given. Yet, as we show in Appendix C.2.4, a version of our model in which advertising is informative about the product’s quality (as in, for instance, Butters (1977), Grossman and Shapiro (1984) or Milgrom and Roberts (1986)), would also feature the demand shifter in equilibrium.

Finally, in our model, advertising is a static choice. Advertising nonetheless embodies a general-
equilibrium effect that critically changes the dynamic incentives of acquiring new product lines through R&D, and it thus has an indirect effect on firm and economic growth. For completeness, in Appendix C.2.1, we show how to accommodate for the “goodwill” accumulation of advertising within our theoretical framework. This does not affect the qualitative results that we obtain in our baseline model.

4.1.4 Production Sector

At any instant, there is an endogenously determined set $\mathcal{F}$ (of measure $F > 0$) of active incumbent producers operating in a monopolistically competitive product market. Each good $j \in [0, 1]$ is produced by a single firm $f \in \mathcal{F}$, and a single firm may own multiple goods. A firm owns a product if it can produce it at a higher intrinsic quality than any other firm. Hence, a typical firm $f$ is summarized by the countable set of products for which it has the leading-edge technology, denoted $\mathcal{J}_f \subseteq [0, 1]$. The number of active product lines owned by firm $f$ is $n \equiv |\mathcal{J}_f| \in \mathbb{Z}_+$. Henceforth, the variable $n$ is referred to as the size of a firm $f$.\footnote{As we shall see, the only way for firms to acquire new product lines is through R&D, and therefore the variable $n$ may be alternatively interpreted as a proxy for the firm’s stock of knowledge.} Finally, the product quality portfolio of firm $f$ is given by $q_f \equiv \{q_j : j \in \mathcal{J}_f\} \in \mathbb{R}^n_+$. Each good variety $j$ is produced with linear technology:

$$y_{jt} = \bar{Q}_t l_{jt}$$

where $l_{jt}$ is labor input and $\bar{Q}_t \equiv \int_0^1 q_{jt} dj$ is the average intrinsic quality in the economy. This linear formulation implies that all good producers face the same marginal cost, $w_t/\bar{Q}_t$. Moreover, by making production scale with aggregate quality, we can ensure that output grows at the same rate as productivity, which is necessary for the existence of a balanced growth path equilibrium.

4.2 Quality Improvements

We assume that R&D-induced intrinsic quality innovations in the production sector are the only engine of growth in the economy. Furthermore, firms can choose to advertise on their currently owned product lines in order to increase demand through the extrinsic quality margin. We turn to both of these types of quality improvements next.
4.2.1 R&D: Improving Intrinsic Quality ($q$)

Firms can invest into R&D to generate the possibility of intrinsically improving their own product lines (so-called *internal* innovation), as well as to creatively build upon the intrinsic quality of those goods that are transferred from other incumbent monopolists through a process of Schumpeterian creative destruction (so-called *external* innovation). Firms expand or contract in the product space on the basis of these innovations.

First, internal R&D is undertaken by incumbent firms in order to technologically advance their existing product lines. To create a Poisson flow rate $z_j \geq 0$ of improving the intrinsic quality of product $j \in J_f$, a firm $f$ must spend a cost of $R_z(z_j)$ units of the final good. This cost function is assumed to be convex in the innovation rate and linear in quality, so that

$$R_z(z_j) = \hat{\chi} q_j z_j^{\hat{\psi}}$$

where $\hat{\chi} > 0$ and $\hat{\psi} > 1$. If successful, internal R&D improves intrinsic quality by a factor of $(1 + \lambda^I) > 1$, so that

$$q_{j,t+\Delta t} = (1 + \lambda^I)q_{jt}$$

if there is only one innovation within an arbitrarily small time interval of size $\Delta t > 0$. The fact that $R_z$ increases with $q_j$ captures that more advanced technologies face higher R&D costs.

Second, external R&D is undertaken by incumbents and potential entrants to obtain technological leadership over products that they do not currently own in their product portfolio. For simplicity, external R&D is assumed to be undirected, so that the successful innovator uniformly draws a good from the set $[0, 1]$.

Incumbents and entrants face, however, different innovation technologies.

On the one hand, entrants (firms with $n = 0$) must incur an expenditure of

$$R_e(x_e) = \nu \bar{Q} x_e$$

units of the final good in order to generate a Poisson flow rate of $x_e$, where $\nu > 0$ is a constant parameter. We assume that there is a measure-one mass of potential entrants, and determine $x_e$ by a free-entry condition.\(^{29}\)

\(^{28}\)By the law of large numbers, this draws a good that is almost-surely not in the innovator’s current portfolio.  
\(^{29}\)By normalizing the mass of potential entrants to unity, and because of linearity in the cost function, $x_e$ denotes in
On the other hand, in order to create a flow Poisson rate of $X_n$, an incumbent firm with $n \geq 1$ product lines must incur an expenditure of $R_x(X_n, n)$ units of the final good. In particular, the cost function is assumed to be convex in the rate of innovation, $X_n$, such that

$$R_x(X_n, n) = \bar{\chi} Q X^\psi_n n^\sigma$$

where $\bar{\chi} > 0$, $\psi > 1$, and $\sigma \leq 0$. If successful, external R&D improves intrinsic quality by a factor of $(1 + \lambda^E) > 1$, so that

$$q_{j,t+\Delta t} = (1 + \lambda^E) q_{jt}$$

for a randomly selected product line $j \in [0, 1]$, if there is only one innovation within an arbitrarily small time interval of size $\Delta t > 0$.

**Discussion**  We define total R&D expenditure by a firm of size $n$ choosing innovation rates ($\{z_j\}, X_n$) by:

$$R_n \equiv \sum_{j \in J} R_z(z_j) + R_x(X_n, n) \quad (3)$$

Slightly abusing terminology, we say that there exist respectively decreasing, constant and increasing returns to scale in R&D if an $n$-product firms finds it respectively more expensive, as expensive and less expensive to grow through innovation by a given rate than $n$ firms of one product each. Since R&D is the sole engine of firm growth in this economy, the degree of returns to scale in $R_n$ determines the predictions of the model regarding growth and R&D intensity across firms of different sizes.

On the one hand, as we shall see, the specification for $R_z$ ensures that $z_j = z$, $\forall j$, because both costs and benefits of internal innovations are linear in quality. This is equivalent as internal R&D scaling proportionally with firm size, which we label as constant returns to scale. Therefore, the degree of returns to scale in the R&D technology operates solely through external innovations.

Accordingly, the fact that a firm has built a stock of knowledge over time ($n \geq 1$) creates heterogeneity in external R&D returns when $\sigma \neq 0$. Absent our advertising margin, it is this scalability which determines whether the model can deliver the empirically observed deviations from Gibrat’s law documents equilibrium both the optimal Poisson rate of entry and the realized mass of entrants. We define the entry rate of the economy, in turn, as $x_e/F$. 

---

19
mented above. With constant returns to scale in innovation costs (i.e, \( \psi + \sigma = 1 \)), for instance, ex-ante technological advantages do not exist with respect to differences in firm size. In other words, in a model without advertising, external innovation investment would in this case scale up one-for-one with added product lines: for any constant \( x > 0 \), a firm of size \( n \) that wishes to create an external R&D intensity \( x \equiv \frac{X_n}{n} \) has the same total expenditure as that of \( n \) firms of size one, or \( R_x(nx, n) = n \cdot R_x(x, 1) \). Consequently, the cost of growing by a given rate is the same for a firm of size \( n \) than for \( n \) firms of size one. This naturally delivers a theoretical version of Gibrat’s law, as in the model of Klette and Kortum (2004): firm value is proportional to size, and thus firm growth is independent of \( n \).

Because this specialization of the model is ill-suited to analyze, among other empirical regularities, the reasons why smaller firms are more innovation-intensive in the data, subsequent work has relaxed the assumption of homogeneity of degree one on returns to innovation. Akcigit and Kerr (2015) extend the Klette and Kortum (2004) framework to incorporate decreasing returns to external R&D (i.e, \( \psi + \sigma > 1 \)). This provides a technological advantage to small firms in terms of R&D: for any constant \( x > 0 \), a firm of size \( n \) that wishes to create a given external R&D intensity \( x \equiv \frac{X_n}{n} \) has a higher total expenditure than that of \( n \) firms of size one, or \( R_x(nx, n) > n \cdot R_x(x, 1) \). As a result, external innovation scales weakly relative to internal innovation, and small firms optimally choose to use external innovations more intensively.

In contrast, we remain agnostic regarding the degree of returns to scale of R&D. In our model with advertising, we are able to qualitatively replicate empirical patterns of the data (e.g., deviations from Gibrat’s law, and decreasing R&D and advertising intensities across firm sizes) when there exist constant, and even increasing, returns to scale in R&D.

Our formulation of advertising effectiveness, which relies on empirical patterns that we draw from the marketing literature, delivers as an equilibrium outcome that smaller firms are more concerned with innovation even when technological advantages in innovation might not be particularly tailored toward them. This is because those firms are the ones that marginally benefit the most in terms of increased advertising efficiency from an increase in the size of their portfolio of products. This leads them to prefer to expand their product portfolios and effectively be relatively more R&D-intensive. In doing so, we obtain that smaller firms invest more in R&D per dollar that they sell, in line with the evidence presented in Section 3. This is true even when

---

\(^{30}\)We will show that we can obtain a quantitative match by calibrating a model with constant returns \((\psi + \sigma = 1)\), but we shall also report results for both decreasing and increasing returns as a robustness check in the comparative statics exercises of Section 4.6.2. Additionally, in Section 7.2, we present a calibration where firms do not have the opportunity to advertise and there are decreasing returns to scale in R&D.
there are constant returns to scale in R&D. This is a novel explanation that highlights the potential effects of firms’ advertising on firm growth which can have significant policy implications especially in terms of R&D subsidies as discussed in Section 7.\footnote{In Appendix C.1, we extend our baseline model with a theory of patent citations following Akcigit and Kerr (2015). We show that our model can deliver two patterns which have sometimes also been interpreted as evidence of technological differences in R&D across firm size, i.e., the fact that smaller firms tend to produce relatively more patents and that those patents tend to be of higher quality.}

### 4.2.2 Advertising: Improving Extrinsic Quality ($\phi$)

Besides building upon the intrinsic quality component of their products, firms can undertake advertising expenditures in order to enhance quality on the extrinsic dimension.

Specifically, for a firm of size $n \geq 1$ with a portfolio of intrinsic qualities $q$, we assume the following function for advertising effectiveness on good $j \in J_f$:

$$\phi(m_j, n) = \tilde{\theta}_j m_j^\zeta n^\eta$$

where $m_j$ is the expenditure into advertising good $j$, with $(\zeta, \eta) \in \mathbb{R}_+^2$, and $\tilde{\theta}_j$ is a good- and firm-specific efficiency component, defined by

$$\tilde{\theta}_j \equiv \theta \frac{q_j}{Q_q} Q^{1-\zeta}$$

where $Q_q \equiv \left( \sum_{q_j \in q} q_j^{1/\alpha} \right)^\alpha$ is a within-firm measure of aggregate intrinsic quality, and $\theta \geq 0$ is the component of advertising efficiency that is constant across time, goods and firms.

**Discussion**  
Equation (4) embodies some of the main effects of advertising identified by the marketing literature reviewed in Section 2 that are necessary within our framework to confirm the four main empirical facts listed in Section 3.

First, everything else equal, the return to advertising a good $j$ is increasing in advertising expenditures on that same good. The parameter $\zeta$ is the elasticity of advertising returns to advertising expenditure. There is ample evidence in Marketing supporting the assumption of diminishing returns to advertising expenditures. For instance, Simon and Arndt (1980), Sutton (1991) and Jones (1995) share the view that advertising response functions are concave.\footnote{Bagwell (2007) (section 2) reviews this literature and concludes that most studies find similar results.} The concavity assumption in the advertising function is also standard in economic models with marketing, e.g. Stigler (1961), Arkolakis
(2010) and Dinlersoz and Yorukoglu (2012). Accordingly, henceforth we assume that $0 < \zeta < 1$.

Second, the marketing return increases in the object $q_j/Q_q$, which is a measure of the relative quality of the good with respect to the aggregate quality within the firm. The latter is expressed in the form of an aggregator across intrinsic quality levels within the firm’s portfolio. In words, advertising is more effective for goods that are intrinsically of relatively higher quality within the firm. This circumstance is identified empirically by the marketing literature reviewed in Section 2, e.g. Archibald, Haulman, and Moody (1983), Caves and Greene (1996), Marquardt and McGann (1975), Rotfeld and Rotzoll (1976), Bagwell (2007) and Kirmani and Rao (2000).

Third, and critically for our mechanism, larger firms have an absolute marketing advantage over smaller firms in that, everything else equal, the return to advertising is greater for higher $n$. We interpret this as there being spillover effects within the firm coming from good-specific advertising expenses: to the extent that conducting advertising increases demand and ultimately sales, owning multiple product lines is beneficial in that it increases revenues over all goods. This effect is meant to capture the value of firm branding, as it is reminiscent of the “umbrella branding” effect discussed in the empirical marketing literature which we reviewed in Section 2, e.g. Lane and Jacobson (1995), Tauber (1981, 1988), Rangaswamy, Burke, and Oliver (1993), Smith and Park (1992), Balachander and Ghose (2003), Erdem (1998) and Erdem and Sun (2002). In the model, $\eta > 0$, so that the fact that the firm owns multiple goods builds a brand value that is attributed evenly across all goods in its portfolio.

**4.3 Entry, Exit, and Pricing**

There is no exogenous exit, and firms move endogenously on the size distribution in a step-wise fashion via the external innovation margin. Creative destruction occurs whenever a good is taken over by a successful external innovator, and a firm exits when it loses its last remaining product.

We assume that a firm has the right to produce a good if it possesses a technological leadership on it, that is, if it can produce the good at the highest *intrinsic* quality among all active firms. The last innovator in each product line then owns the leading patent and has monopolistic power until it is displaced by another firm. Once a firm makes an innovation, it acquires a perpetual patent on it. However, this patent does not preclude other firms from investing into research to improve the intrinsic quality of the product. As it is standard in the literature, we assume that the new innovator is then able to price the old incumbent out of the market. Specifically, we assume that firms entering with
new innovations can charge the unconstrained monopoly price, as opposed to having to price-compete with the current incumbent.\footnote{This could also be obtained as an equilibrium result if firms in a given product line compete \`a la Bertrand and if there exists an arbitrarily small cost of posting a price as in Akcigit and Kerr (2015). In this case, only the technological leader chooses to pay the posting price.}

4.4 Resource Constraints

The economy is closed and there is no government, so GDP equals aggregate consumption plus aggregate investment. The latter is split between aggregate R&D expenditure by entrants and incumbents, denoted $Z_t$, and aggregate advertising expenditure by incumbents, denoted $M_t$. The resource constraint at time $t$ is

$$C_t + Z_t + M_t \leq Y_t \quad (5)$$

In the labor market, labor demand comes from intermediate goods producers, so that

$$\int_0^1 l_j dj \leq 1$$

for all $t \in \mathbb{R}_+$.

4.5 Equilibrium

In this section we derive the Markov Perfect Equilibrium of the economy at any given time. Later on, we specialize the equilibrium to a balanced-growth path (BGP) in which all aggregates grow at a constant and positive rate, $g$.

**Consumer’s Problem** Taking initial wealth $A_0$ as given, the representative consumer chooses a path for consumption to maximize utility subject to the flow budget constraint and the no-Ponzi condition $\lim_{t \to +\infty} e^{\int_0^t \rho ds} A_t \geq 0$. The optimality condition yields the standard Euler equation for consumption:

$$\frac{C_t}{C_t} = r_t - \rho \quad (6)$$

Intermediate-good firms are owned by the household, so the value of household wealth is equal to
the value of corporate assets. Namely, $A_t = \int_{\mathcal{F}} V(q_f) df$, where $V(q_f)$ denotes the net present value of the whole future expected stream of profits for a monopolist $f$ who owns the (intrinsic) quality portfolio $q_f$ at time $t$. In equilibrium, the following transversality condition holds:

$$\lim_{t \to +\infty} \left[ e^{-\int_0^t r_s ds} \int_{\mathcal{F}} V_t(q_f) df \right] = 0 \quad (7)$$

**Final Good Firm’s Problem**  The representative final good producer solves the following problem, taking qualities $\{\tilde{q}_j : j \in [0, 1]\}$ and input prices $\{p_j : j \in [0, 1]\}$ as given:

$$\max_{y_{jt}, j \in [0, 1]} \left\{ Y_t - \int_0^1 p_{jt} y_{jt} dj \right\}$$

s.t. $Y_t = \frac{1}{1 - \beta} \int_0^1 \tilde{q}_{jt} y_{jt}^{1 - \beta} dj$

This leads to:

$$p_{jt} = \left( \frac{\tilde{q}_{jt}}{y_{jt}} \right)^{\beta} \quad (8)$$

Recalling that $\tilde{q}_{jt} = q_{jt}(1 + d_{jt})$, Equation (8) says that the inverse demand function for goods is iso-elastic, and $\beta$ is the price-elasticity. This equation makes apparent that advertising, by increasing total quality ($\tilde{q}_{jt}$) through its extrinsic component, effectively impacts consumer preferences, because it works as a perfect demand shifter that alters consumption decisions.34

**Incumbent Firm’s Problem**  A monopolist chooses labor, quantities, prices, R&D and advertising expenditures over each good in its portfolio in order to maximize the present discounted value of the total future stream of profits. Our set-up allows us to break this problem into a static part, in which the firm sets price, quantity and advertising expenditures over the goods that it currently owns, and a dynamic part, in which R&D decisions are made.

Before the R&D and advertising choices, the static maximization problem of monopolist $j \in [0, 1]$ holding the patent for the leading-edge intrinsic quality of good $j$ is:

34In Appendix C.2.5, we present an extension of the model in which advertising not only acts as a shifter in the demand function, but can also affect its price-elasticity.
\[ \pi(\tilde{q}_{jt}) = \max_{y_{jt}} \left\{ p_{jt} y_{jt} - w_{jt} l_{jt} \right\} \]

\[ \text{s.t. } y_{jt} = \tilde{Q}_t l_{jt} \text{ and } p_{jt} = \tilde{q}_{jt} y_{jt}^{-\beta} \]

where \( \pi \) stands for operating profits before advertising and R&D expenditures. The optimality condition implies \( p_{jt} = \left( \frac{1}{1-\beta} \right) \frac{w_{jt}}{\tilde{Q}_t} \), meaning that all monopolists set the same price every period, or \( p_{jt} = p_t, \forall j \). This price is the optimal unconstrained monopoly solution: a constant markup over the marginal cost. Using labor market clearing, \( \int_0^1 l_{jt} = 1 \), we can find the market-clearing wage:

\[ w_t = (1 - \beta) [\tilde{Q}_t + \Phi_t]^{\beta} \tilde{Q}_t^{1-\beta} \quad (9) \]

where \( \Phi_t \equiv \int_0^1 \phi_{jt} dj \) denotes aggregate extrinsic quality. The monopoly price can thus be written as:

\[ p_t = \left( \frac{\tilde{Q}_t + \Phi_t}{\tilde{Q}_t} \right)^{\beta} \quad (10) \]

Aggregate output is:

\[ Y_t = \frac{1 - \beta}{\beta} [\tilde{Q}_t + \Phi_t]^{\beta} \tilde{Q}_t^{1-\beta} \quad (11) \]

Flow operating profits before advertising costs can be written as:

\[ \pi_{jt} = \tilde{\pi}_t \tilde{q}_{jt} \quad (12) \]

where

\[ \tilde{\pi}_t = \beta \left( \frac{\tilde{Q}_t}{\tilde{Q}_t + \Phi_t} \right)^{1-\beta} \quad (13) \]

is the component of profits that is constant across goods. From Equation (13), we can notice that the profit of a firm is a decreasing function of the overall investment in advertising in the economy through \( \Phi_t \). Equation (12) shows that any difference in flow profits across goods must be due to differences in either intrinsic or extrinsic quality, or both. Specifically, two goods of the same intrinsic quality can earn different profits if the firm chooses to advertise one more than the other. Therefore,
R&D and advertising represent two non-mutually exclusive ways to increase profits.

We next describe how the good-specific advertising decision is optimally made by the firm.

**Advertising Choices**  Henceforth, we drop time subscripts unless otherwise needed. When choosing advertising expenditures \( m_j \), a firm of size \( n \) and portfolio of (intrinsic) qualities \( q \in \mathbb{R}_+^n \) solves the static problem:

\[
\pi_{adv} \equiv \max_{\{m_j, j \in J\}} \sum_{q_j \in q} \left[ \tilde{\pi}(q_j + \phi(m_j, n)) - m_j \right] \\
\text{s.t. } \phi(m_j, n) = \theta \frac{q_j}{\left( \sum_{q_{j'} \in q} q_{j'}^{1/\alpha} \right)^{1 - \zeta} \bar{Q}^{1 - \zeta} n^\eta} \tag{14}
\]

where \( \pi_{adv} \) denotes flow profits after the advertising decision (i.e., net of advertising costs). We make the assumption that \( \alpha = 1 - \zeta \) in order to reduce the dimensionality of the parameter space and be able to find closed-form expressions.\(^{35}\) The optimality condition gives good-specific advertising expenditures of

\[
m_j = \left( \zeta \theta \tilde{\pi} \right)^{1 - \zeta} \frac{q_j^{1/(1 - \zeta)}}{\sum_{j' \in q} q_{j'}^{1/(1 - \zeta)} \bar{Q}^{1 - \zeta} n^\eta} \tag{15}
\]

and thus firm-level advertising expenditures of

\[
M_n \equiv \sum_{q_j \in q} m_j = \left( \zeta \theta \tilde{\pi} \right)^{1 - \zeta} \bar{Q}^{1 - \zeta} n^\eta \tag{16}
\]

for a firm of size \( n \). We define optimal advertising intensity by \( M_n/n \). In our empirical analysis, after controlling for firm observables and time fixed-effects, we have found that (i) advertising expenditures are increasing in size (delivered directly by the fact that \( \eta > 0 \)), and that (ii) advertising intensity is decreasing in size (Fact #3). To satisfy the second requirement, \( M_n \) must be concave in \( n \), which leads to the following parametric restriction: \( \eta + \zeta < 1 \).

This explicitly implies that \( \eta < 1 \) (assuming diminishing returns to advertising i.e. \( \zeta < 1 \)). Since this parameter controls for the degree of size spillovers coming from advertising (recall Equation (4)), we obtain that, in order to replicate the decreasing advertising intensity, the spillover effect from size

\(^{35}\)Our results would not be qualitatively affected by the choice of a different value of \( \alpha \) because the main mechanism regarding advertising and R&D intensity in our model works through firm size \( n \).
must be marginally stronger for smaller firms. Namely, the gain in advertising return for an \( n \)-to-(\( n+1 \)) transition is higher when \( n \) is smaller.

As a consequence, smaller firms will be relatively more concerned with expanding to new product markets, and they will choose a relatively higher external innovation intensity in equilibrium. This mechanism implies that the interaction between R&D and advertising can generate a decreasing R&D intensity with firm size (Fact #2) and hence a deviation from Gibrat’s law (Fact #1). Our model can then qualitatively generate Facts 1 to 3 even in the presence of non-decreasing returns to scale in R&D.

We find that total extrinsic quality at the good level is:

\[
\phi_j = \bar{\gamma} \frac{q_j^{1/(1-\zeta)}}{\sum_{j'} q_{j'}^{1/(1-\zeta)}} Q_n^{1-\zeta} 
\]  

(17)

and that total firm-level extrinsic quality is:

\[
\Phi_n = \sum_{q_j \in q_f} \phi_j = \bar{\gamma} Q_n^{1-\zeta} 
\]  

(18)

for a firm of size \( n \), where \( \bar{\gamma} \equiv \theta(\zeta \bar{\pi})^{1-\zeta} \). Importantly, the last two equations show that the empirical facts that motivated our specification of the advertising structure in Equation (4) feature in the optimal choice of firms as long as we impose \( \eta + \zeta < 1 \). At the good level (Equation (17)), we see that (i) goods benefit more from advertising when the firm is larger (i.e, \( \phi_j \) increases in \( n \)); (ii) through the advertising spillover effect, a given good benefits relatively more from an innovation-driven \( n \)-to-(\( n+1 \)) transition when it is produced by a smaller firm (i.e, \( \phi_j \) is concave in \( n \)); and (iii) for given size, relatively higher-quality goods within the firm benefit more from marketing (i.e, \( \phi_j \) increases in \( q_j \)). Such equilibrium size advantages exist at the firm level as well (Equation (18)): for a given portfolio of goods, firm-level extrinsic quality is higher for larger firms, but when advancing the quality portfolio from \( n \) to \( n+1 \) goods through external innovation, smaller firms experience a relatively higher increase in extrinsic quality (as the marginal firm-level increase in extrinsic quality, given by \( (\Phi_{n+1} - \Phi_n) \), is a decreasing function of \( n \)).

Combining returns and costs, we have that the overall static profits of the firm are

\[
\pi^{adv} = \bar{\pi} \sum_{q_j \in q} q_j + \bar{\gamma} Q_n^{1-\zeta} 
\]  

(19)
where $\gamma \equiv \tilde{\pi}\tilde{\gamma} - (\theta\tilde{\pi}\zeta)^{\frac{1}{1-\zeta}}$ is constant across goods and firms. Equation (19) is the static value of firm-level flow operating profits net of advertising expenditures. It has two components. The first one grows linearly with the aggregate intrinsic quality within the firm, with a factor of proportionality that does not vary across firms (although it typically may grow with the economy). The second one is invariant to the firm’s intrinsic quality portfolio, but is increasing and concave in firm size. This component comes from investment in advertising and the extrinsic margin of product quality. While the first term is standard in models of endogenous firm growth with scalable returns, the second term is new and critically alters the dynamic incentives to conduct research vis-à-vis advertising in the way that we have described above.

**BGP Equilibrium Characterization** A Balanced Growth Path equilibrium is defined as an equilibrium allocation in which output all aggregate variables grow at a constant rate, denoted by $g > 0$. To show the existence of such an equilibrium, we make use of the fact that economy-wide extrinsic quality $\bar{\Phi}$ in fact grows at the same rate as that of aggregate intrinsic quality, $\bar{Q}$. In this sense, R&D remains the only engine of growth in the model, as in the original formulation by Klette and Kortum (2004).

Formally, we have that:

$$\bar{\Phi} = \Phi^*\bar{Q}$$

where $\Phi^*$ is a constant, both across time and firms. Under this result, it is clear from Equation (11) that, if a BGP exists where output grows at the rate $g$, then $g = \frac{\bar{Q}}{\bar{q}_t}$. Moreover, we can note from Equation (13) that $\tilde{\pi}$ is time-invariant, and therefore flow operating profits before advertising costs can be written as $\pi_{jt} = \pi\tilde{q}_{jt}$ in BGP, where $\pi = \beta(1 + \Phi^*)^{\beta-1}$ is a constant. In turn, $\gamma$ is then constant, and $\Phi$ becomes a linear function of $\bar{Q}$, from Equation (18). By construction, we can express total aggregate extrinsic quality on the BGP as:

$$\bar{\Phi} = \sum_{n=1}^{+\infty} F\mu_n \Phi_n$$

where $\mu_n$ is the invariant share of size-$n$ firms (which we derive explicitly below), such that $\mu_n \in [0, 1]$ and $\sum_{n=1}^{+\infty} \mu_n = 1$. Combining $\bar{\Phi} = \Phi^*\bar{Q}$ with the last equation and the equilibrium firm-level extrinsic
quality found in (18), we then obtain that $\Phi^*$ is the fixed point of the following expression:

$$
\Phi^* = \theta^{1-\zeta} \left( \frac{\zeta \beta}{1 + \Phi^*)^{1-\beta}} \right)^{1-\zeta} \sum_{n=1}^{+\infty} F \mu_n n^{\frac{\eta}{1-\zeta}}
$$

Equation (20) delivers a unique solution $\Phi^* > 0$.\(^{36}\) In our numerical algorithm described below, we exploit this fixed-point to solve for the unique BGP of the economy by iterating over $\Phi^*$. A direct consequence of the result is that aggregate marketing expenditures, $M$, grow at the rate $g$, since

$$
M = \sum_{n=1}^{+\infty} F \mu_n M_n, \quad \text{and } M_n \text{ has been shown to be linear in } \bar{Q} \text{ (Equation (16)). In particular, we}
$$

obtain

$$
M = \sum_{n=1}^{+\infty} F \mu_n (\zeta \theta \pi)^{1-\zeta} n^{\frac{\eta}{1-\zeta}} \bar{Q}
$$

and thus $\dot{M}/M = g$. From the resource constraint in (5), if total R&D expenditures $Z$ grow at the rate of $\bar{Q}$ (a result whose proof we relegate to the end of this section), so does aggregate consumption, $C$. From Equation (6), we then have that $g = r_t - \rho$, and therefore $r_t = g + \rho, \forall t$.

**Value Functions** Denote by $\tau$ the endogenous rate of creative destruction along the BGP, i.e. the aggregate rate of external innovation coming from both entrants and incumbents. Denote by $x_n$ the external R&D intensity of a firm of size $n$, i.e. $x_n \equiv X_n/n$, where $X_n$ is the Poisson rate of external innovation. Taking $(r, \tau, g)$ as given, a firm $f$ with a product portfolio with $1 \leq n = |q_f|$ goods chooses external R&D intensity $x_n$, internal R&D intensities $\{z_j : j \in J_f\}$, and advertising expenditures $\{m_j : j \in J_f\}$ to maximize firm value $V_n(q_f),^{37}$ written in the Hamilton-Jacobi-Bellman form (see Appendix B.1 for a derivation):

$$
r V_n(q_f) = \max_{x_n,\{z_j,m_j\}} \left\{ \sum_{q_j \in q_f} \left[ \tilde{\pi}(q_j + \phi_j) - \tilde{x} z_j q_j - m_j + z_j \left( V_n(q_f \setminus \{q_j\} \cup \{q_j(1 + \lambda^I)\}) - V_n(q_f) \right) \right] + \tau \left( V_{n-1}(q_f \setminus \{q_j\}) - V_n(q_f) \right) + nx_n \left( E_j V_{n+1}(q_f \cup \{q_j(1 + \lambda E)\}) - V_n(q_f) \right) - \tilde{x}^{\sigma + \tilde{x} \pi} \bar{Q} \right\} + \dot{V}_n(q_f)
$$

\(^{36}\)This is because the right-hand side is continuous, strictly positive and decreasing in $\Phi^*$, since $\zeta, \beta \in (0,1)$.

\(^{37}\)Note that the optimal advertising investment $(m_j)$ was derived in Equation (15).
subject to Equation (4), where $\cup_+$ and $\backslash_-$ are multiset union and difference operators. \(^{38}\) The right-hand side of Equation (22) is composed of multiple terms. The first three terms in the first line are good-specific flow operating profits net of internal R&D costs and advertising expenditures. The fourth term in the first line is the change in value due to the internal improvement over a currently held good. This event occurs at a Poisson flow rate of $z_j$, and it increases the intrinsic quality of the good by a factor of $(1 + \lambda^I) > 1$. The first term in the second line is the change in the firm’s value when losing a good $j$ to another incumbent or an outside entrant through creative destruction, an event that occurs at the equilibrium Poisson rate $\tau$. In this case, the firm becomes of size $n - 1$. The second term on this line is the change in value for the firm when it successfully acquires a new product line through external innovation and transitions to $n + 1$, which occurs at the Poisson rate $n x_n$ (recall that here $x_n$ denotes the intensity rate). The third line includes the flow resources spent in external R&D, and the change in firm value that occurs due to growth at the economy-wide level (the term $\dot{V}_n(q_f)$).

As for entrants (firms with no products), their value function is given by the following equation (see Appendix B for a derivation):

$$
rv_0 = \max_{x_e > 0} \{ x_e \left[ E_j V_1(q_j (1 + \lambda^E)) - V_0 \right] - \nu x_e \bar{Q} \} + \dot{V}_0
$$

(23)

where $V_0$ denotes firm value at $(n, q_f) = (0, \emptyset)$. Entrants are choosing the rate of entry $x_e$, in which event they randomly draw a good $j \in [0, 1]$ and improve upon intrinsic quality by a factor of $(1 + \lambda^E) > 1$. If there is a positive mass of entrants in equilibrium, the free-entry condition reads

$$
V_0 = 0
$$

which implies that $\dot{V}_0 = 0$. Our first proposition shows that the growth rate of the economy is measured as the combined contributions of internal and external innovators, which advance aggregate productivity by $\lambda^I$ and $\lambda^E$, respectively. To be precise, the proposition uses that $z_j = z, \forall j \in [0, 1]$, a result that we shall prove shortly.

**Proposition 1 (Growth rate)** The growth rate of the economy in the BGP equilibrium is given by

$$
g = \tau \lambda^E + z \lambda^I
$$

(24)

\(^{38}\)These operators are defined by $\{a, b\} \cup_+ \{b\} = \{a, b, b\}$, and $\{a, b, b\} \backslash_- \{b\} = \{a, b\}$, and they are needed because the set $q_f$ may include more than one instance of the same element.
Moreover, since innovation rates are independent, the rate of creative destruction $\tau$ is given by the aggregate Poisson rate of external innovation coming from entrants and incumbents. While $x_e$ denotes both the realized Poisson rate of entry and the total equilibrium mass of entrants, we must take into account that there exist different shares of incumbents according to their size, and a total mass $F$ of incumbents at any point in time. Therefore,

$$\tau = x_e + \sum_{n=1}^{+\infty} F \mu_n x_n \quad (25)$$

In turn, we can find a closed-form expression for $\mu_n$, the time-invariant share of firms of size $n \in \mathbb{Z}_+$ along the BGP:

**Proposition 2 (Invariant firm-size distribution)** The invariant firm-size distribution is given by

$$\mu_n = \frac{x_e \prod_{i=1}^{n-1} x_i}{F n \tau^n}; \quad \forall n \geq 1 \quad (26)$$

with $\mu_n \in [0, 1], \forall n \geq 1$, and $\sum_{n=1}^{+\infty} \mu_n = 1$.

**Proof.** See Appendix B.3.

We are now ready to solve for the value function and find optimal R&D intensities. To this purpose, we guess a functional form of the type:

$$V_n(q_f) = \Gamma \sum_{q_j \in q_f} q_j + \Upsilon_n \bar{Q}$$

for which we must verify the value of the constant coefficient $\Gamma \in \mathbb{R}_+$ and the sequence $\{\Upsilon_n\}_{n=1}^{+\infty}$. This guess is informed by the expression for the flow value of holding a given product, seen in (19). Namely, a firm derives value from the intrinsic quality of the products it is selling, independently of its size (the first term, whose coefficient $\Gamma$ is constant across firm types), but also from the option value of marketing its goods in order to increase demand and revenue (the second term, whose coefficient $\Upsilon_n$ depends on firm type).\(^{39}\)

The following proposition summarizes the solution:

\(^{39}\)When $\tilde{\psi} + \sigma \neq 1$, $\Upsilon_{n+1} - \Upsilon_n$ also captures the dependence of external R&D on firm size.
Proposition 3 (Value functions) Assuming that there is positive entry in equilibrium \( x_e > 0 \), the value for a firm of size \( n \geq 1 \) is 
\[ V_n(q_f) = \Gamma \sum_{q_j \in q_f} q_j + \Upsilon_n \bar{Q}, \]
where
\[ \Gamma = \frac{\nu - \Upsilon_1}{1 + \lambda^E} \]  \hspace{1cm} (27)
and \( \Upsilon_n \), for \( n \geq 1 \), is the solution to the second-order difference equation
\[ \Upsilon_{n+1} - \Upsilon_n + \Gamma (1 + \lambda^E) = \bar{\vartheta} \left( \rho \Upsilon_n n^{\frac{\sigma}{\psi - 1}} - (\Upsilon_{n-1} - \Upsilon_n) \tau n^\sigma + \Upsilon_{n+1} - \Upsilon_n \right) \]  \hspace{1cm} (28)
where \( \bar{\vartheta} \equiv \widetilde{\psi} \left( \frac{\widetilde{\chi}}{(\psi - 1)^{\psi - 1}} \right)^\frac{1}{\psi} \) is a parameter, with boundary condition \( \Upsilon_0 = 0 \).

Proof. See Appendix B.4.

In the appendix we show that we obtain the following optimal R&D intensities:

\[ z_j = \left( \frac{\lambda^I (\nu - \Upsilon_1)}{\psi \bar{X} (1 + \lambda^E)} \right)^{\frac{1}{\psi - 1}} \]  \hspace{1cm} (29)
\[ x_n = n^{\frac{1 - \sigma - \bar{\psi}}{\psi - 1}} \left( \frac{\nu - \Upsilon_1 + \Upsilon_{n+1} - \Upsilon_n}{\psi \bar{X}} \right)^{\frac{1}{\psi - 1}} \]  \hspace{1cm} (30)

for internal and external innovations, respectively, and all \( n \geq 1 \). First, we confirm that for internal improvements the optimal innovation effort is independent of the quality of the product line \( z_j = z \), since both costs and returns are linear in \( q_j \). On the other hand, \( n \) affects external innovation intensities through two distinct channels: (i) the degree of scalability in external R&D technology \( (\sigma + \bar{\psi}) \); (ii) the degree of decreasing returns in advertising through the branding effect (i.e., the values for \( \zeta \) and \( \eta < 1 - \zeta \)). The former channel features both explicitly in the exponent of the first multiplicative term of (30), and implicitly within the \( \Upsilon_n \) sequence in (28). The advertising channel, however, only features indirectly through \( \Upsilon_n \). Thus, we can obtain the desired inverse dependence of innovation expenditures to size as long as the advertising channel scales weakly enough to compensate any strong scalability coming from the R&D technology itself. This allows for the possibility of balanced growth in equilibrium even when there exist increasing returns to scale in external R&D.

For the sake of intuition, let us focus on the case in which there are constant returns in external R&D (namely, internal and external scale proportionally with size), which shall be the case of analysis...
in our baseline calibration. In this case, $\sigma + \tilde{\psi} = 1$ and Equation (28) reads

$$\Upsilon_{n+1} - \Upsilon_n + \Gamma(1 + \lambda^E) = \tilde{\vartheta} \left( \frac{\rho \Upsilon_n}{n} - (\Upsilon_{n-1} - \Upsilon_n)\tau - \frac{\gamma n^{\eta/1-\zeta}}{n} \right)^{\frac{\tilde{\psi}-1}{\tilde{\psi}}}$$

The left side of the equation is the total change in value that is associated with an $n$-to-$(n+1)$ transition, and therefore it measures the marginal gain for a firm of size $n$ of increasing in size through successful external innovation. In the absence of the advertising channel ($\gamma = 0$) and under constant returns to scale in R&D, the model is equivalent to that of Klette and Kortum (2004). In this case, $\Upsilon_n$ becomes $\Upsilon_n = n \Upsilon$ for some constant $\Upsilon > 0$. This implies that the gain in value from acquiring one additional good (i.e., $\Upsilon_{n+1} - \Upsilon_n$) is constant across firm size and we obtain the theoretical results that (i) external R&D investment scales perfectly with size, and (ii) consequently firm growth is independent of size (Gibrat’s law). When $\theta > 0$, however, deviations from Gibrat’s law can occur even when $\sigma + \tilde{\psi} = 1$: small firms might still be more innovation-intensive than large firms because they benefit marginally more in terms of advertising efficiency gains from the new product lines that they acquire through external R&D (thanks to advertising spillovers across different goods).

In particular, gains depend non-linearly on $n$. As a small firm (low $n$) accumulates more products, marginal gains to external R&D decrease because the marginal gain in advertising efficiency is decreasing. This means that the absolute value of external R&D for a firm of size $n$ is increasing and concave, and thus the marginal gain of acquiring an additional product line for a firm of size $n$, given by $(\Upsilon_{n+1} - \Upsilon_n)$, is decreasing and convex. In the limit as $n$ grows larger, the marginal gain converges to a positive constant.

**Equilibrium Definition** To close the optimality characterization and define an equilibrium, it remains to show two conditions that have been stated without proof in the above discussion.

First, the assumption of positive entry ($x_e > 0$) stated in Proposition 3 relies on the equilibrium condition $\nu < \Upsilon_1$ (so that $\Gamma > 0$). Intuitively, this condition says that the cost of entry $\nu$ is lower than the gain, coming from the acquisition of one product line. This condition is always checked ex-post in our computation.

Second, from (5) and the result that $g = \dot{Q}/\bar{Q} = \dot{C}/C = \dot{M}/M$, we need to check that aggregate

---

40: We relegate the analysis with decreasing and increasing returns to the comparative statics of Section 4.6.2.

41: The left hand side is actually the expected gain from acquiring an additional good normalized by average intrinsic quality in the economy. The term $\Gamma(1 + \lambda^E)$ is the change in value coming from the increase in the aggregate intrinsic quality of the firm’s portfolio when acquiring one more product. Since this change in value is common across sizes, it does not play a direct role on innovation incentives across $n$. 

33
R&D expenditure $Z$ grows at the rate $g$, as well. Using the cost functions for R&D, we have that:

$$Z = \int_0^1 R_z(z_j)dz + \sum_{n=1}^{\infty} F_\mu_n R_x(X_n, n) + R_e(x_e)$$

$$= \left(\hat{\chi}_z^{\hat{\psi}} + \sum_{n=1}^{\infty} F_\mu_n \hat{\chi}_n \hat{\mu}_n^{\hat{\psi}} + \nu x_e \right) \bar{Q}$$

(31)

The right-hand side includes aggregate R&D expenditures by incumbents attempting internal innovations, incumbents attempting external innovations, and outside firms attempting to enter into the market through external R&D, respectively. Indeed, we get that $Z$ is linear in $\bar{Q}$, and hence $\dot{Z}/Z = g$.

Aggregate consumption is then computed as the residual $C = Y - Z - M$.

We are now ready to define a BGP equilibrium:

**Definition 1 (BGP Equilibrium)**  A Balanced Growth Path Equilibrium is, for all $q_j \geq 0$, $j \in [0, 1]$, $n \in \mathbb{Z}_+$, $\bar{Q}$ and $t \in \mathbb{R}_+$, allocations $y_j$, $l_j$ and $m_j$; extrinsic quality $\phi_j$; aggregates $Y$, $C$, $Z$, $M$, $A$; a mass of incumbents $F$; a constant $\Gamma$ and a sequence $\{\Upsilon_n\}_{n=1}^{+\infty}$; a firm size distribution $\mu_n$; prices $w$, $p$, and $r$; and rates $g$, $z$, $x_n$, $x_e$, and $\tau$; such that:

- Given prices, final good producers maximize profit;
- $y_j$ (input quantity) and $p$ (input price) solve the intermediate sector problem and satisfy (10);
- $m_j$ (expenditure) and $\phi_j$ (effectiveness) solve the advertising problem and satisfy (15) and (17), respectively;
- Innovation flows $z$ (internal for incumbents) and $x_n$ (external for incumbents) solve the innovation problem and satisfy (29) and (30);
- $\mu_n$ (the invariant measure of firms) satisfies (26);
- $x_e$ (entry flow) and $F$ (mass of incumbent firms) solve the entry problem and satisfy the free-entry condition $V_0 = 0$ and the restriction $\sum_n \mu_n = 1$;
- $\Gamma$ (the value of a firm’s intrinsic-quality portfolio) and $\Upsilon_n$ (the value of owning $n$ products) solve the value function in (22), and satisfy (27) and (28), respectively;
- $g$ (growth rate) and $\tau$ (creative destruction rate) satisfy (24) and (25), respectively;
- Aggregates $Y$ (output), $M$ (advertising expenditure), and $Z$ (R&D investment) satisfy (11), (21), and (31), respectively, and $C$ (consumption) is found residually via (5);
• $A$ (household wealth) satisfies $A = \sum_n F_n V_n$, and the transversality condition (7);
• $r$ (interest rate) satisfies $r = g + \rho$;
• $w$ (wage) clears the labor market.

4.6 Computation

Before turning to the estimation of the model in Section 5, we discuss its computational implementation and present a few comparative statics results for a given parametrization in order to illustrate our main advertising-innovation trade-off.

4.6.1 The BGP Algorithm

To compute the equilibrium variables in the BGP, we loop over the fixed point of $\Phi_t$ by using $\Phi_t = \Phi^* \tilde{Q}_t$, with $\Phi^*$ from Equation (20). In light of the above discussion, we check for convergence by imposing that the value of firm size, $\Upsilon_n$, converges to a line as $n$ grows larger. We truncate the size space at some large $N \in \mathbb{N}$. The following describes the steps of the algorithm:

1. Guess a number $\Phi^* > 0$.
2. Compute $\Upsilon_1$ by imposing $\Upsilon_2 = 2\Upsilon_1$ in Equation (28) at $n = 1$.
   
   (a) Compute $g$ from (24), $\tau$ implied by (25), $\Gamma$ from (27), and $z$ from (29).
   
   (b) Compute $\{\Upsilon_n\}_{n=2}^{+\infty}$ using (28) forward from $\Upsilon_1$ and $\Upsilon_0 = 0$, $\{x_n\}_{n=1}^{+\infty}$ from (30), and $F/x_e$ using that $\sum_{n=1}^N \mu_n = 1$, where $\mu_n$ comes from (26).
3. Verify convergence of firm size distribution. If there is no convergence by iteration $k \in \mathbb{N}$, go back to step 2 with the new guess for $\Upsilon_1$ to be the solution to (28) at $n = 1$ when $\Upsilon_2 = 2\Upsilon_1 - \varepsilon_k$, for a small $\varepsilon_k > 0$.
4. Compute $\Phi^*$ as the solution to (20), and compare it to the initial guess. If it is too far, go back to step 1 using this solution as the new guess.

In step 2, we compute the maximum value for $\Upsilon_1$ such that $\Upsilon_n$ can be weakly concave (i.e., $(\Upsilon_{n+1} - \Upsilon_n)$ decreasing). In particular, we force $\Upsilon_n$ to be straight line from $n = 1$ to $n = 2$ (note $\Upsilon_2 - \Upsilon_1 = \Upsilon_1 - \Upsilon_0$). If $\mu_n$ does not converge, it must be because $(\Upsilon_n - \Upsilon_{n-1})$ has not settled to a flat line as $n$ has approached $N$, which means that the guess for $\Upsilon_1$ was incorrect. Then, we iterate on new guesses.

\footnote{We further verified that there exists only one value of $\Upsilon_1$ for which we obtain convergence for a given $\Phi^*$.}
for $\Upsilon_1$ to allow for more concavity on the $\Upsilon_n$ sequence (indeed, note that in any iteration $k \geq 1$ we always start the $\Upsilon_n$ sequence at a $\Upsilon_1$ such that $\Upsilon_2 - \Upsilon_1 < \Upsilon_1 - \Upsilon_0$). In step 3, we in turn bisect the new guess by a factor of ten on each new iteration, i.e. $\varepsilon_{k+1} = \varepsilon_k / 10$. Finally, in step 4, we drop the old $\Phi^*$ guess in case of no convergence, and use the resulting fixed-point as the new guess.

4.6.2 Comparative Statics on BGP (Numerical Example)

To illustrate the qualitative features of the model, this section presents a set of comparative statics exercises on outcomes of the balanced growth path equilibrium with respect to the following parameters: advertising efficiency ($\theta$) and the degree of returns to scale in innovation ($\tilde{\psi} + \sigma$). Throughout, we set $\lambda^E = \lambda^I$ so that neither type of innovation is more radical than the other. The purpose of this section is threefold: (i) we show how our model can generate decreasing R&D intensity with firm size under constant returns to scale in R&D, (ii) we confirm that the model can potentially deliver both substitutability and complementarity between R&D and advertising (our calibration will identify the relevant case in Section 5) and (iii) we show that our model can be solved with increasing returns to scale in R&D technology. All the parameter values are set for expositional purposes, and we relegate the discussion of the calibration to Section 5.1.

Advertising Efficiency ($\theta$) Figure 2 shows, as a function of firm size ($n$), BGP results for the marginal gain of acquiring new product lines (i.e. $\Upsilon_{n+1} - \Upsilon_n$) on the left panel, the expenditure intensity on external R&D (i.e. $R_x/n$) on the middle panel, and the expenditure intensity on advertising (i.e. $M_n/n$) on the right panel for different values of $\theta$. Figure 3 shows that our framework can generate an invariant firm-size distribution that resembles a Pareto.

First, in the absence of advertising ($\theta = 0$), we return to a Klette and Kortum (2004) economy in which innovation incentives are constant in size. The gain of acquiring new products is proportional to size, or $\Upsilon_n = n \Upsilon$, and therefore expenditures are constant in the cross-section of firms. When we introduce a motive for advertising ($\theta > 0$), firms start investing in advertising. Moreover, because smaller firms benefit marginally more from expanding their product portfolios ($\eta < 1$), they invest more in R&D relative to larger firms. Consequently, the firm-size distribution in Figure 3 is more right-skewed for any $\theta > 0$ than the one delivered by the special case of Klette and Kortum (2004): fewer firms achieve large scales because, as they keep growing, they become increasingly less concerned

---

43 The parameter values used in this example are: $\rho = 0.02$, $\tilde{\psi} = \tilde{\psi} = 1 - \sigma = 2$, $\lambda^E = \lambda^I = 0.05$, $\tilde{\chi} = 0.1$, $\chi = 10$, $\nu = 2$, $\beta = 0.2$, $\eta = 0.5$ and $\zeta = 0.1$.  

36
with expanding their size further through R&D.

As $\theta > 0$ increases further (with all other parameters fixed), advertising becomes more efficient for all levels of firm size, and the aforementioned effects are reinforced. Moreover, since advertising now delivers higher returns, in the margin small firms now benefit even more than they did before, relative to larger firms, as reflected by the fact that the marginal gain in firm value, $(\Upsilon_{n+1} - \Upsilon_n)$, becomes steeper as $\theta$ increases. As a consequence, expenditures decrease faster with size ($R_x/n$ and $M_n/n$ are also steeper). In other words, the advertising spillover effect tilts dynamic R&D incentives toward smaller firms. In this example, we can notice that R&D and advertising are substitutes since larger values of $\theta$ are associated with lower investment in R&D and larger advertising expenditures. Next, we provide an example in which R&D and advertising are complements.

**R&D and advertising as complements** In this section, we show that our model can also generate complementarity between R&D and advertising. Figure 4 shows the change in external R&D across firm size for different levels of advertising efficiency as well as total incumbent R&D and advertising as
In this example, R&D and advertising are complements. An increase in advertising efficiency is associated with larger investment in both advertising and R&D. As we shall see, whether R&D and advertising are complements or substitutes plays a crucial role in determining the growth and welfare effect of advertising. In Section 5.1, we calibrate our model to determine which case is the relevant one in the data.

Returns to Scale in Innovation  Finally, we show that the result that small firms are more innovation-intensive does not hinge on the relative degree of scalability in the returns of different types of innovation technologies. In Figure 5 we show three solutions of the advertising model, with decreasing, constant, and increasing returns to scale in external R&D, respectively.

Decreasing returns lower innovation incentives for all sizes with respect to constant returns, but make it even more profitable for small firms to invest more intensely into R&D (i.e, the $R_x/n$ line steepens). Symmetrically, having increasing returns in innovation increases optimal intensity for all $n$. Interestingly, it is still the case that smaller firms optimally choose to invest relatively more, even if the marginal benefit is lower than before (i.e, the $R_x/n$ line flattens but it is still decreasing).

The possibility of solving the model when there exist increasing returns in R&D is an interesting feature of our model, given that the special case of Klette and Kortum (2004) (i.e, $\theta = 0$) has no

---

The parameter values used in this example are: $\rho = 0.02$, $\hat{\psi} = \bar{\psi} = 2$, $\sigma = -0.85$, $\lambda^E = \lambda^I = 0.05$, $\bar{\chi} = 1$, $\bar{\chi} = 5$, $\nu = 2$, $\beta = 0.2$, $\eta = 0.9$ and $\zeta = 0.3$. 

---

Figure 3: Firm size distribution in BGP, for different values of advertising efficiency ($\theta$).
Figure 4: External R&D intensity (left panel) and R&D expenditure as a share of GDP (right panel) in BGP, for different values of advertising efficiency ($\theta$).

Note: External R&D expenditure intensity is normalized by aggregate intrinsic quality, $\bar{Q}$.

Figure 5: Comparative statics for $\sigma + \tilde{\psi} = 1.05 > 1$ (decreasing returns), $\sigma + \tilde{\psi} = 1$ (constant returns), and $\sigma + \tilde{\psi} = 0.95 < 1$ (increasing returns).
solution when $\sigma + \tilde{\psi} < 1$.

5 Quantitative Analysis

In this section, we calibrate the model combining the micro data used in Section 3 with aggregate moments for the U.S. in the period 1980 - 2015. We use the same sample from Compustat as in Section 3 and seek to match empirical regression coefficients with model-implied slopes from simulation-based regressions (indirect inference). This calibration strategy is fairly standard in the endogenous growth and firm dynamics literature (e.g. Lentz and Mortensen (2008)). The novelties are that we use coefficients on advertising regressions to calibrate our theory of firm growth and, most importantly, that we are able to match the slopes of both firm growth and R&D intensity regressions without relying on decreasing returns to scale in the R&D technology.

5.1 Calibration Strategy

There are 13 parameters to identify: $(\rho, \lambda^E, \lambda^I, \bar{\chi}, \tilde{\psi}, \hat{\psi}, \sigma, \nu, \beta, \eta, \zeta, \theta)$. Some of these parameters are externally identified, while some others are internally calibrated via our indirect inference approach.

In our baseline calibration, we impose constant returns to scale in external innovation ($\sigma = 1 - \tilde{\psi}$) and assume that no innovation is more radical than the other ($\lambda^E = \lambda^I \equiv \lambda$). Once again, we do this to isolate the effects of our advertising-R&D interaction on firm growth in order to show that the spillover effect on advertising is not only able to quantitatively match advertising intensity slopes, but also the observed deviations from Gibrat’s law and R&D intensity across sizes.

External Identification The parameters $(\rho, \hat{\psi}, \tilde{\psi}, \beta, \zeta)$ are externally calibrated. We set $\rho = 0.02$, which approximately corresponds to a discount factor of 97% annually. We impose $\hat{\psi} = \tilde{\psi} = 2$, following Akcigit and Kerr (2015) and prior empirical literature estimating the cost curvature of different types of R&D. To find a value for $\beta$, we use (8), (10), and (12) to obtain $\beta = \frac{\int_0^1 \pi_{jt} \frac{dj}{dj}}{\int_0^1 p_{jt} y_{jt} \frac{dj}{dj}}$. Using our Compustat sample described in Section 3, we calculate the corresponding ratio of average operating income before depreciation to average sales, and find $\beta = 0.1645$. Finally, we set $\zeta = 0.1$ which is in line with estimations of the elasticity of sales to advertising expenditures found in the empirical marketing literature.\footnote{We take the number from Tellis (2009), who in turn estimates it as the average elasticity over a sample of about 260 estimates gathered from prior studies. $\zeta$ is approximately equal to the elasticity of sales to advertising expenditures. To see
**Internal Identification**  We are left with the parameters \((\lambda, \bar{\chi}, \hat{\chi}, \nu, \eta, \theta)\), which are determined internally. In particular, we solve and simulate the economy and find parameter values that match model moments and simulation-implied regression coefficients to those observed in the data and presented in tables in Section 3. The invariant equilibrium is computed using the algorithm described in Section 4.6.1, whereas the simulation of the model uses 1,000 firms and discretizes time to \(T = 100\) periods of length \(\Delta t = 0.01\).\(^{46}\)

We find values for the 6 internally identified parameters by targeting 7 moments for our period of interest. We target the aggregate firm entry rate from the Business Dynamics Statistics (BDS) data of the U.S. Census Bureau,\(^{47}\) the long-run average growth rate of per capita GDP in the U.S., and the average ratios of R&D-to-sales and R&D-to-advertising expenditures in our sample. In addition, we run simulation-based regressions to target the 4 empirical coefficients of interest that relate to the empirical facts presented in Section 3: the Gibrat’s law coefficient (column (3) in Table 1), the R&D intensity and advertising intensity coefficients (column (3) in Tables 2 and 3, respectively), and the R&D-to-advertising coefficient (column (3) in Table 4).\(^{48}\) To obtain these simulation-based results, we compute the corresponding coefficients via OLS and take the time-series average of the corresponding coefficients.\(^{49}\) Throughout, we use sales as our measure of size.\(^{50}\) Our selection criterion is a simple unweighted minimum absolute-distance estimator.\(^{51}\)

Table 5 shows the full set of calibrated parameter values. Table 6 shows the results for our targeted moments, compared to the values in the data. We do well on all targeted moments. Particularly, all four simulation-based slopes have the correct sign and magnitude. Even though there are constant returns to scale in the R&D cost function, the spillover effect in advertising, which is set to target this, firstly note from Equation (4) that

\[
\frac{\partial \log d_j}{\partial \log m_j} = \zeta,
\]

where recall that \(d_j = \phi_j/q_j\). Moreover, per-product sales \(p_jy_j\) are linear in \(\tilde{q}_j = q_j(1 + d_j)\), so

\[
\frac{\partial \log (p_jy_j)}{\partial \log m_j} = \frac{\partial \log (p_jy_j)}{\partial \log (1+d_j)} \approx \frac{\partial \log d_j}{\partial \log m_j} = \zeta,
\]

where we have used \(\frac{\partial \log (p_jy_j)}{\partial \log (1+d_j)} = 1\), and the approximation is valid in BGP as the economy grows large.\(^{46}\)

All firms are assumed to start the same, with qualities \(q_0 = 1\). In the event of ties, namely multiple successful firms drawing the same good over which to perform an external innovation, we assign equal probabilities to each firms and draw the new monopolist from the corresponding discrete uniform distribution.\(^{47}\)

Note that the coefficient on the R&D-to-advertising substitution across sizes can be obtained from combining the coefficients on R&D and advertising intensities, which means that we are in fact targeting three of the four coefficients.\(^{48}\) To allow for convergence of the simulation, we ignore the first 20 periods in our calculations.\(^{49}\) In the model, the measure of firm size is number of products. Our data has no information on this dimension. However, we rely on a number of empirical studies finding a positive correlation between the two. For instance, Scherer (1983) and Katila and Ahuja (2002) show that the introduction of new products positively correlates with sales growth and market value. More directly, Plehn-Dujowich (2013) shows that there is a high positive correlation between the number of products of a firm and its employment and revenues.\(^{50}\)

In particular, we find the minimum of \(\sum_{m=1}^{k} \frac{|\text{Model}(m) - \text{Data}(m)|}{\text{Data}(m)}\), where \(k\) is the number of moments.\(^{51}\)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>0.02</td>
<td>Time discount rate</td>
<td>Standard</td>
</tr>
<tr>
<td>$\psi$</td>
<td>2</td>
<td>Elasticity of R&amp;D</td>
<td>Akcigit and Kerr (2015)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1645</td>
<td>Profitability ratio</td>
<td>Compustat</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>0.1</td>
<td>Sales-ADV elasticity</td>
<td>Tellis (2009)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$1 - \psi$</td>
<td>Constant RTS in R&amp;D</td>
<td>Assumption</td>
</tr>
</tbody>
</table>

**Externally identified**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>0.0143</td>
<td>Innovation step</td>
<td>.</td>
</tr>
<tr>
<td>$\hat{\chi}$</td>
<td>0.0017</td>
<td>Internal R&amp;D scale param.</td>
<td>.</td>
</tr>
<tr>
<td>$\tilde{\chi}$</td>
<td>0.6256</td>
<td>External R&amp;D scale param.</td>
<td>.</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.7206</td>
<td>Entrant’s R&amp;D scale param.</td>
<td>.</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.8527</td>
<td>ADV spillover effect</td>
<td>.</td>
</tr>
<tr>
<td>$\theta$</td>
<td>1.1022</td>
<td>ADV efficiency</td>
<td>.</td>
</tr>
</tbody>
</table>

**Internally identified**

Table 5: Full set of calibrated parameters in baseline estimation (model with advertising and constant returns to scale in R&D).

decreasing advertising intensity with size observed in the data, is also able to correctly predict both the decreasing R&D intensity and the deviation from constant firm growth. This implies that the observed decreasing R&D intensity in firm size observed in the data is not necessarily evidence for the existence of decreasing returns to scale in R&D. Our calibration shows that it could simply be the result of the interaction between R&D and advertising even when there exist constant returns to scale in R&D. This can have potential policy implications on the optimal design of R&D subsidies (see Section 7).

Figures 6 and 7 show a graphical depiction of the calibrated economy. Figure 6 plots sales in logs against the growth rate of sales, where each circle represents a simulated firm. We see that, on average, larger firms tend to grow slower: as in the data, a 1% increase in sales decreases the growth rate of firms by 0.0326%, on average. The left panel of Figure 7 in turn shows the distribution of sales, normalized by its mean, in the simulated economy. On the right panel, we observe that there is a very close link between the number of products that a firm has and its (normalized) level of sales.

### 5.2 Growth Effects of Advertising

In this section, we study the effect of a decrease in advertising cost (an increase in $\theta$) on the growth rate of the economy. As we showed in Section 4.6.2, R&D and advertising can be either complements or substitutes depending on parameter values. We now show that in our calibrated model R&D and
### Aggregate moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Model</th>
<th>Data</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average growth rate $g$</td>
<td>0.02</td>
<td>0.02</td>
<td>Standard</td>
</tr>
<tr>
<td>Firm entry rate $x_e/F$</td>
<td>0.101</td>
<td>0.098</td>
<td>BDS</td>
</tr>
<tr>
<td>Average R&amp;D-Sales ratio (\sum_n F\mu_n R_n/(py))</td>
<td>0.153</td>
<td>0.102</td>
<td>Compustat</td>
</tr>
<tr>
<td>Average R&amp;D-ADV ratio (\sum_n F\mu_n R_n/M_n)</td>
<td>24.15</td>
<td>26.40</td>
<td>Compustat</td>
</tr>
</tbody>
</table>

### Regression coefficients

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>Data</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gibrat’s coefficient</td>
<td>$\beta_{\Delta\text{sales/sales}}$</td>
<td>-0.0326</td>
<td>-0.0325</td>
</tr>
<tr>
<td>R&amp;D intensity coefficient</td>
<td>$\beta_{\text{rd/sales}}$</td>
<td>-0.1030</td>
<td>-0.1035</td>
</tr>
<tr>
<td>ADV intensity coefficient</td>
<td>$\beta_{\text{adv/sales}}$</td>
<td>-0.0353</td>
<td>-0.0317</td>
</tr>
<tr>
<td>R&amp;D/ADV coefficient</td>
<td>$\beta_{\text{rd/adv}}$</td>
<td>-0.0677</td>
<td>-0.0719</td>
</tr>
</tbody>
</table>

Table 6: Targeted moments: model vs. data. Model with advertising and constant returns to scale in R&D.

Notes: Sales is used as the measure of firm size. \(R_n\) stands for the total R&D expenditures of a firm of size \(n\), defined in Equation (3). Similarly, \(M_n\) is total advertising expenditures at the firm level, defined in Equation (16).

![Figure 6: Firm growth rates on (normalized) sales levels in the calibrated economy.](image)
advertising act as substitutes at the firm level. This implies that more efficient advertising can have a detrimental effect on growth by decreasing the rate of innovation in the economy.

Interestingly, even when R&D and advertising are substitutes, the effects of advertising efficiency on growth is ambiguous. To see this, we can decompose the growth rate of the economy by combining (24) and (25) to obtain:

\[ g = \underbrace{\lambda E x_e}_\text{Entrants} + \underbrace{z \lambda I}_\text{Incumbents doing internal} + \lambda E \sum_{n=1}^{+\infty} F \mu_n x_n \text{ Incumbents doing external} \]  

(32)

The three additive terms in \( g \) correspond to the contribution to growth by entrants, incumbents performing internal improvements, and incumbents performing external innovations, respectively. In addition, there is a compositional effect coming from the change in the distribution of firm. Since firms of different size have different R&D investment behavior, this compositional channel might also affect the rate of growth of the economy.

In Figures 8 and 9, we plot these different effects. First, because advertising is now more efficient, the value of becoming an incumbent firm increases and there is more entry (higher \( x_e \)). The increased R&D investment by entrants has a positive effect on innovation and economic growth. It also increases the rate of creative destruction which modifies the incentive to perform internal R&D. Second, a higher creative destruction rate implies that firms are more likely to lose one of their products by being
displaced by another firm. As a result, firms have smaller incentives to perform internal R&D. This decrease in internal R&D investment depresses economic growth. Third, incumbent firms also decrease their external R&D intensity when advertising becomes more efficient. This decrease is even more pronounced for larger firms. Finally, there is a compositional effect operating through changes in the firm size distribution. As a result of higher entry rate and lower external innovation by incumbent, the firm size distribution shifts to the left (Figure 9). This places more mass on the part of the firm distribution that is most R&D intensive. This compositional channel thus tends to foster growth.

![Graphs showing the effects of advertising on economic growth](image)

**Figure 8:** Growth decomposition, for different values of advertising efficiency ($\theta$), in the calibrated economy. The calibrated value of $\theta$ is marked by the dashed vertical line. The growth rate is $g = \tau \lambda^E + z \lambda^I$, where $\tau = x_c + \sum_{n=1}^{\infty} \mu_n n x_n$ is the creative destruction rate. The entry rate is measured by $x_e/F$.

Overall, the decrease in internal and external R&D investment more than offset the increase in entry rate and compositional change resulting in a reduction of growth when advertising becomes more efficient. In particular, we find that growth would increase by 0.64 percentage point (from 2% to 2.64%) if the advertising sector was shut down completely within the calibrated economy (namely, if $\theta$ changed from its calibrated value, marked by the dashed vertical line in the plots, to $\theta = 0$).

### 5.3 Welfare Effects of Advertising

There has been a long-lasting debate in the economic literature about the welfare implications of advertising. This literature has studied the effects of advertising on welfare through mainly two
potential channels: informative advertising and taste-shifting advertising.

Models of informative advertising are theories in which advertising can be used to remove frictions in markets with imperfect information by providing relevant information about a product quality (or simply about the existence of the product). Such approach can be found for instance in Nelson (1974), Butters (1977), and Grossman and Shapiro (1984). In that context, advertising could be welfare improving. On the other hand, Becker and Murphy (1993) argue that advertising could simply be used to alter consumer preferences by creating favorable associations to the good that is being advertised. Kihlstrom and Riordan (1984) and Milgrom and Roberts (1986) further argue that expenditures in uninformative advertising could still signal a product quality and hence provide ex-post information.

Another strand of the literature considers advertising as a pure taste shifter by manipulating consumer preferences (see among others Dixit and Norman (1978), Becker and Murphy (1993) and Benhabib and Bisin (2002, 2011)). Molinari and Turino (2009) consider the case of purely combative advertising, in which advertising does not affect consumer preferences in equilibrium but firms nevertheless advertise to maintain their market shares. In this case, advertising expenditures are a pure waste of resources. When advertising directly affects tastes, a relevant question is whether welfare should be compared using ex-ante or ex-post preferences. In addition, advertising could distort prices and monopoly power, in which case advertising could be welfare decreasing even when measured by ex-post preferences (Dixit and Norman (1978), Benhabib and Bisin (2002)). In this section, we argue

\[ \theta = 1.1 \text{ (calib.)} \]
\[ \theta = 0.3 \]

**Figure 9:** Change in firm size distribution as a function of advertising efficiency ($\theta$).
that one should consider yet another side effect of advertising when quantifying its welfare implications, namely its effect on R&D investment, innovation and economic growth.

To compute ex-post welfare along a BGP, we can rewrite the household welfare function in equilibrium as:

$$U(C_0, g) = \frac{\ln (C_0)}{\rho} + \frac{g}{\rho^2}$$

where $C_0$ is proportional to $\hat{Q}_0 = 1$ (a normalization). To study the effect of advertising, we compare ex-post welfare in the calibrated economy across different advertising efficiency levels ($\theta$). Adopting an ex-post view means that we take into account the level effects that advertising has on perceived consumption. To show that this modeling choice is not critical, in Appendix C.2.3 we present a version of the model in which advertising is a wasteful activity, in the sense that it does not increase the value of consumption in equilibrium. Making welfare comparisons in that case would be equivalent to comparing utility using ex-ante preferences. Of course, with ex-ante preferences, only the growth effect is operative and welfare is decreasing in advertising efficiency ($\theta$).

The first term in the ex-post welfare decomposition accounts for a “level effect” of advertising: by conducting advertising, firms expand demand contemporaneously, which increases the perceived utility derived from aggregate consumption. This is because advertising ultimately constitutes a demand shifter for the firm by effectively changing consumer’s preferences, as seen in Equation (8). Additionally, as we have argued, advertising critically shapes the dynamic incentives to conduct R&D and, therefore, has an indirect effect on growth. This is captured by the second term, which we label the “growth effect” of advertising.

While advertising unambiguously increases welfare through the level effect, its impact through the growth effect is negative in our calibrated model. Figure 10 plots welfare and its level-growth decomposition as a function of $\theta$. Once again, the level effect is unambiguously positive and, as we have just discussed, advertising decreases growth in our calibration, so the final effect is ambiguous. However, we find that welfare is reduced because the growth effects dominates. According to this result, therefore, a policy scheme in which the advertising sector is taxed can be welfare-improving for the economy. We analyze this implication in Section 7.1.

In sum, the fact that R&D and advertising crowd each other out is critical for the analysis. In Section 6.2, we will show that this substitution is present in the data, as well.
Our results add to the existing debate about advertising and welfare by showing that, even in a model in which advertising effectively increases utility (through a “level effect” in our model but potentially from information in alternative formulations), one should consider the potential side effects of cheaper advertising. We show that one such side effect is its impact on economic growth in the long run coming from the substitution between R&D and advertising.

6 Validation Exercises

Our validation exercises are made of two parts. First, we compare untargeted moments in our model and in the data. In particular, we look at the relationship between firm size and the variance of our main firm-level observables, namely firm growth, and R&D and advertising intensities. We also look at other correlations and the degree of persistence of R&D and advertising investment. Second, we test the R&D-advertising substitution that is predicted by our calibrated model and which is at the core of our growth and welfare results. We use exogenous variations in R&D cost arising from changes in R&D tax treatment across U.S. states and investigate the response of advertising investment to those variations.
6.1 Untargeted Moments

In this section, we compare some untargeted moments from the calibrated model to the data. The first set of moments that we consider is the variance of firm growth, R&D and advertising intensities across firm size. It is a well-known stylized fact in the literature that not only the average but also the variance of firm growth rates decreases with firm size (see among others Hymer and Pashigian (1962), Evans (1987) and Klette and Kortum (2004)). In Figure 11, we compare the evolution of these variances in the model and in the data.

![Figure 11: Standard deviations (% with respect to 1st quintile): Model vs Data.](image)

*Notes:* Firms are ranked in size quintiles according to their normalized level of sales (sales as a ratio of average sales in the same year). R&D and advertising intensities are measured as the ratio of total R&D and advertising expenditures to total sales within each group. The standard deviation of firm growth rates in the first quintile is normalized to one. Standard deviations in other quintiles are measured as a share of the first quintile.

Our model fits the relationship between the variance of firm growth and firm size well. In addition, we also report the standard deviation of R&D and advertising intensity with firm size. Figure 11 shows that these standard deviations are also decreasing with firm size in the data. Our model predicts a decrease in the variance in R&D and advertising intensity which is qualitatively and quantitatively in line with the data (the only exception being the increase in R&D intensity variance form the fourth to the fifth quintile in the model).

In Table 7, we compare additional untargeted moments. The correlation coefficient between R&D intensity and firm growth as well as between advertising intensity and firm growth predicted by the
model are broadly in line with those in the data. Our model also predicts high levels of serial correlation in R&D and advertising intensity within firms which are comparable to their empirical counterparts. The ratio of R&D to advertising expenditures is also very persistent both in the model and in the data.

<table>
<thead>
<tr>
<th></th>
<th>DATA</th>
<th>MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr(R&amp;D intensity, firm growth)</td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td>corr(ADV intensity, firm growth)</td>
<td>0.10</td>
<td>0.22</td>
</tr>
<tr>
<td>autocorr(R&amp;D intensity)</td>
<td>0.92</td>
<td>0.89</td>
</tr>
<tr>
<td>autocorr(ADV intensity)</td>
<td>0.88</td>
<td>0.76</td>
</tr>
<tr>
<td>autocorr(R&amp;D/ADV ratio)</td>
<td>0.92</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table 7: Correlation and Autocorrelation Coefficients: Model vs Data.
Notes: R&D and advertising intensity in logarithm.

6.2 An Empirical Investigation of the Substitution between R&D and Advertising

While our theoretical model could potentially generate a complementarity between R&D and advertising, the calibrated version of the model suggests substitutability. Fact #4, presented in Table 4, provided empirical evidence for the substitution of advertising for R&D as firms grow larger. Besides Fact #4, our calibrated model also predicts that a decrease in the price of R&D (and hence in its relative price) would lead to a decrease in advertising intensity, ceteris paribus. Indeed, in Section 5.3 we showed that exogenously increasing the relative price of advertising (that is, decreasing $\theta$) generates an increase in the growth rate of the economy through a substitution between advertising and R&D at the firm level. This crowding-out effect is strong enough to overturn the compositional effects operating along the extensive margin, and is the key mechanism behind the policy implications of Section 7.2. In this section, we provide evidence that the crowding out between advertising and R&D is present in the data as well.

Event To test for the existence of substitution between R&D and advertising expenditure, we use exogenous variation in the cost of R&D over time and across U.S. states. Since the early 1980s, U.S. states started implementing tax incentives for innovation. Starting with the state of Minnesota in 1982, several U.S. states introduced R&D tax incentives to stimulate innovation. These tax incentives are typically in the form of tax credits, corresponding to a certain share of R&D expenditures, which can be deducted from state corporate income tax owned by the firm. The inception date of the tax varies
across states. The tax credit rate also varies across states and over time within each state. This allows us to obtain exogenous variation in the relative price of R&D over time and across states. In 2009, 32 states offered such credits with rates varying between 2% and 20%.

In 1981, an R&D tax credit was also implemented at the federal level. It has since been extended every year with the exception of 1995. The federal tax credit rate was initially set to 25%, and later reduced to 20% from 1986 onwards. This federal credit offers another source of exogenous variation in the cost of R&D that can be used to test for substitution between advertising and innovation. However, it shows little time variation after its introduction and does not vary across states. We shall only use it for computing a measure of the R&D user costs.

**Data** We use annual data from Compustat on R&D, advertising, sales, and our measures of firm age and financial constraints from Section 3, for the period 1950 - 2009. We also retrieve data on the relevant state for tax credit from Compustat. We focus our analysis on firms reporting positive advertising and R&D expenditures. Data on tax credit rates are obtained from Wilson (2009) and Falato and Sim (2014).

To provide evidence for substitution between R&D and advertising, we need to show that a decrease in the relative price of R&D (through an increase in R&D tax credits) leads to a significant decrease in advertising intensity. We provide two sets of empirical evidence supporting the existence of such substitution. First, we look at the relationship between the overall advertising intensity at the state level and exogenous changes in the relative cost of R&D from changes in the tax treatment of R&D expenditures. Second, we perform a similar analysis at the firm level.

**State-level Results** First, let us consider advertising intensity at the state level. Our measure of advertising intensity is the ratio of advertising expenditures within the state divided by the overall level of sales in the state. If R&D and advertising are substitutes, we expect the advertising intensity of a state to decrease when the R&D tax credit in the given state increases, *ceteris paribus*. We perform the following regressions:

\[
\log \left( \frac{Adv_{jt}}{Sales_{jt}} \right) = \alpha_0 + \beta_1 \, Tax \, Credit_{jt} + \beta_2 X_{jt} + \alpha_j + \alpha_t + u_{jt} \tag{34}
\]

for state \( j \) at time \( t \), where \( \alpha_t \) and \( \alpha_j \) control for time and state fixed effects respectively, \( Tax \, Credit_{jt} \)

---

52 We use data for all 50 U.S. states as well as for the District of Columbia.
53 When states have several credit brackets, we use the top marginal rate in our empirical analysis.
is a measure of R&D tax credit and $X_{jt}$ is a vector of control variables including sales in the state, the average age of firms, the average level of equity financing in the state, state corporate tax rates as well as lagged advertising intensity to control for potential serial correlation in advertising expenditures.

Results are reported in Table 8. Column (1) reports the results from regressing advertising intensity on a dummy variable which takes value one if the state has a tax credit implemented at time $t$. The results show a weakly negative correlation between the existence of an R&D tax credit and the level of advertising intensity in the state, conforming with our mechanism. Using this specification, however, ignores differences in the tax credit rate between states and over time. To amend this issue, columns (2) to (5) use different measures of tax credit rate and R&D relative cost in order to take into account the magnitude of the credit.

Column (2) shows a significantly negative relationship between advertising intensity and the relative price of R&D using changes in the statutory tax credit across states and time. In particular, a one-percentage-point increase in the statutory state R&D tax credit is associated with a 1.057% decrease in the ratio of advertising expenditures to sales at the state level. In column (3), we use an alternative measure of the tax credit rate which takes into account the fact that the tax credit is itself subject to corporate taxation in some states. We compute the tax-adjusted credit rate for state $j$ at time $t$ as:

$$ Tax \ Adj. \ Rate_{jt} = Statutory \ Credit \ Rate_{jt} \times (1 - s_{jt} \times Tax \ Rate_{jt}) $$

where $s_{jt}$ is the share of the R&D credit which is subject to corporate taxation. When the credit is taxed, the credit rate is now not only influenced by differences in the statutory credit rate over time and across states, but also by changes in corporate taxes.

Column (4), in turn, uses an alternative measure of the marginal effective R&D tax credit, proposed by Wilson (2009) and available until 2006. This measure acknowledges the different definitions of the R&D expenditures which can lead to a tax credit as well as the horizon over which the tax credit is calculated. In some states, all R&D expenditures can lead to a tax credit, while some other states offer a credit only to R&D expenditures above a certain base level. This threshold can in turn be a moving average of past R&D expenditures. For such states, the moving-average base is usually computed as the product of firm R&D-to-sales ratio over the $n$ previous periods, times current sales. For a firm with R&D expenditures above the base level, the marginal effective tax credit rate ($m_{t}$) is computed as:
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>log ($\frac{\text{Adv}}{\text{Sales}}$)</td>
<td>log ($\frac{\text{Adv}}{\text{Sales}}$)</td>
<td>log ($\frac{\text{Adv}}{\text{Sales}}$)</td>
<td>log ($\frac{\text{Adv}}{\text{Sales}}$)</td>
<td>log ($\frac{\text{Adv}}{\text{Sales}}$)</td>
</tr>
<tr>
<td><strong>Tax credit dummy</strong></td>
<td>-0.125*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0627)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>State credit rate</strong></td>
<td>-1.057**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.464)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tax-adjusted state rate</strong></td>
<td>-1.077**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.466)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Effective state rate</strong></td>
<td>-1.053*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.578)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>R&amp;D user cost</strong></td>
<td>1.148**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.537)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>log(Sales)</strong></td>
<td>-0.244**</td>
<td>-0.237**</td>
<td>-0.237**</td>
<td>-0.244**</td>
<td>-0.246**</td>
</tr>
<tr>
<td></td>
<td>(0.101)</td>
<td>(0.0976)</td>
<td>(0.0976)</td>
<td>(0.110)</td>
<td>(0.111)</td>
</tr>
<tr>
<td><strong>Firm Age</strong></td>
<td>0.0168</td>
<td>0.0175</td>
<td>0.0175</td>
<td>0.0178</td>
<td>0.0160</td>
</tr>
<tr>
<td></td>
<td>(0.0167)</td>
<td>(0.0165)</td>
<td>(0.0165)</td>
<td>(0.0200)</td>
<td>(0.0201)</td>
</tr>
<tr>
<td><strong>Fin. Const.</strong></td>
<td>0.00982*</td>
<td>0.00986*</td>
<td>0.00985*</td>
<td>0.00999*</td>
<td>0.0103*</td>
</tr>
<tr>
<td></td>
<td>(0.00511)</td>
<td>(0.00509)</td>
<td>(0.00509)</td>
<td>(0.00556)</td>
<td>(0.00554)</td>
</tr>
<tr>
<td><strong>State Tax</strong></td>
<td>0.768</td>
<td>0.661</td>
<td>0.666</td>
<td>0.357</td>
<td>-0.742</td>
</tr>
<tr>
<td></td>
<td>(1.541)</td>
<td>(1.608)</td>
<td>(1.608)</td>
<td>(1.686)</td>
<td>(1.754)</td>
</tr>
<tr>
<td><strong>log($\frac{\text{Adv}_0}{\text{Sales}_0}$)</strong></td>
<td>0.724***</td>
<td>0.724***</td>
<td>0.724***</td>
<td>0.707***</td>
<td>0.702***</td>
</tr>
<tr>
<td></td>
<td>(0.0518)</td>
<td>(0.0508)</td>
<td>(0.0508)</td>
<td>(0.0558)</td>
<td>(0.0570)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>-1.937***</td>
<td>-1.944***</td>
<td>-1.943***</td>
<td>-2.051***</td>
<td>-3.904***</td>
</tr>
<tr>
<td></td>
<td>(0.615)</td>
<td>(0.609)</td>
<td>(0.610)</td>
<td>(0.649)</td>
<td>(1.111)</td>
</tr>
</tbody>
</table>

| Time FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| State FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 2084 | 2084 | 2084 | 1933 | 1908 |
| $R^2$ | 0.84 | 0.84 | 0.84 | 0.84 | 0.83 |

**Table 8:** Effect of R&D subsidy on advertising intensity at the state level.

**Notes:** Data from Compustat from 1950 to 2009, Wilson (2009) and Falato and Sim (2014). Age is measured as the elapsed time since the first observation in the data. Our measure of financial constraints is sale minus purchases of common and preferred stock, divided by sales. Age and financial constraint are averaged at the state level. Standard errors are clustered at the state level (in parentheses). Significance level: * 10%; ** 5%; *** 1%.
\[ m_{jt} = \text{Statutory Credit Rate}_{jt} \times (1 - s_{jt} \times \tau^e_{jt}) \times \left( 1 - \frac{1}{n} \sum_{k=1}^{n} (1 + r_{t+k})^{-k} \right) \]

where \( r \) is the real interest rate, \( n \) is the number of periods over which the moving-average base is calculated, and \( \tau^e_{jt} \) is the effective marginal tax rate which takes into account the fact that, in some states, taxes paid to the state can be deducted from federal taxable income, and vice-versa.\(^{54}\)

For states with base definition based on the moving-average of past R&D activity, every dollar spent on R&D today decreases the amount of R&D that qualifies for a tax credit in the future and hence reduces the effective marginal tax credit rate. In some states such as New York and Connecticut, all R&D expenditure qualifies for a tax credit (i.e., there is no moving-average definition of the base level). In this case, the marginal effective credit rate is equal to the (tax-adjusted) statutory credit rate. As shown in column (4) of Table 8, using the marginal effective subsidy rate confirms the negative relationship between advertising intensity and the cost of R&D.

Finally, we use one additional state-level measure of the cost of R&D: the R&D user cost. This measure was extended from Hall and Jorgenson (1967) to R&D investment by Bloom, Griffith, and Van Reenen (2002), and computed at the U.S. state and federal levels by Wilson (2009). In particular, the user cost of R&D in state \( j \) is given by:

\[ \text{R&D User Cost}_{jt} = \frac{1 - v(m_{jt} + m_{ft}) - z(\tau^e_{jt} + \tau^e_{ft})}{1 - (\tau^e_{jt} + \tau^e_{ft})} \left( r_t + \delta \right) \]

where the \( f \) subscript stands for federal, \( v \) is the share of R&D expenditures that qualifies for preferential tax treatment, \( z \) is the present discounted value of tax depreciation allowances, and \( \delta \) is the depreciation rate of R&D capital.\(^{55}\) Higher levels of R&D tax credits are associated with lower user cost of R&D. Column (5) of Table 8 in turn shows that higher R&D user costs are also associated with higher levels of advertising intensity at the state level, as expected if R&D and advertising are substitutes.

**Firm-level Results** Overall, our regression results in Table 8 support the prediction of our calibrated model regarding substitution between R&D and advertising. We now show that this substitution also holds when we look at how advertising expenditures at the firm level change following an exogenous variation in the cost of R&D. We use the same measures of R&D tax credit and user cost as in

---

\(^{54}\)This rate also takes into account the fact that R&D tax credit are subject to corporate taxation in some states.

\(^{55}\)Wilson (2009) sets \( v = 0.5, z = 0.525 \) and \( \delta = 0.15 \).
the previous analysis. We perform the following regressions:

$$
\log \left( \frac{Adv_{ijzt}}{Sales_{ijzt}} \right) = \alpha_0 + \beta_1 \text{Tax Credit}_{jt} + \beta_2 X_{ij,t} + \alpha_t + \alpha_j + \alpha_z + u_{ijzt}
$$

(35)

for firm $i$ in state $j$ and industry $z$ at time $t$, where $\alpha_t$, $\alpha_z$ and $\alpha_j$ control for time, industry and state fixed effects respectively, Tax Credit$_{jt}$ is a measure of R&D tax credit and $X_{ij,t}$ is a vector of control variables including sales, the age of the firm, the level of equity financing, state corporate tax rates as well as lagged advertising intensity to control for potential serial correlation in advertising expenditures.

Regression results are reported in Table 9. We obtain a significant negative relationship between the state R&D tax credit and advertising intensity at the firm level. In particular, a one-percentage-point increase in the statutory state tax credit is associated with a decrease in advertising intensity of around 0.4%. Similarly, higher R&D user cost are associated with higher advertising intensity.

Taken together, the results of Tables 8 and 9 suggest that R&D and advertising are substitutes in our sample, in line with the predictions of our calibrated model. They provide empirical support to the mechanism through which advertising has significant implications in terms of overall R&D expenditures, innovation, economic growth and welfare.

7 Policy Implications

We now present potential policy implications of our model. First, since we have showed that advertising can be detrimental to growth and welfare, we analyze the possibility of a tax on advertising. Second, we have showed that the observed decreasing R&D intensity with firm size could potentially have two sources: the interaction between R&D and advertising (as in our baseline calibration) and decreasing returns to scale to R&D. Our second policy exercise therefore compares the effectiveness of R&D subsidies under our baseline calibration versus a world with technological differences in R&D production function.

---

56 The idea of taxing advertising is not new and has been for instance discussed in Solow (1968).
Table 9: Effect of R&D subsidy on advertising intensity at the firm level

<table>
<thead>
<tr>
<th></th>
<th>(1) log(Adv/Sales)</th>
<th>(2) log(Adv/Sales)</th>
<th>(3) log(Adv/Sales)</th>
<th>(4) log(Adv/Sales)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State credit rate</td>
<td>-0.428**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.168)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tax-adjusted state rate</td>
<td>-0.414**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.175)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective state rate</td>
<td></td>
<td>-0.454**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.207)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R&amp;D user cost</td>
<td>0.380*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.218)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(Sales)</td>
<td>0.00138</td>
<td>0.00133</td>
<td>-0.00127</td>
<td>-0.00133</td>
</tr>
<tr>
<td></td>
<td>(0.00320)</td>
<td>(0.00320)</td>
<td>(0.00348)</td>
<td>(0.00348)</td>
</tr>
<tr>
<td>Firm Age</td>
<td>0.000916*</td>
<td>0.000925*</td>
<td>0.000815</td>
<td>0.000835</td>
</tr>
<tr>
<td></td>
<td>(0.000531)</td>
<td>(0.000531)</td>
<td>(0.000592)</td>
<td>(0.000592)</td>
</tr>
<tr>
<td>Fin. Const.</td>
<td>0.000500**</td>
<td>0.000500**</td>
<td>0.000464**</td>
<td>0.000465**</td>
</tr>
<tr>
<td></td>
<td>(0.000232)</td>
<td>(0.000232)</td>
<td>(0.000213)</td>
<td>(0.000213)</td>
</tr>
<tr>
<td>State tax</td>
<td>-1.824***</td>
<td>-1.813***</td>
<td>-1.878***</td>
<td>-2.243***</td>
</tr>
<tr>
<td></td>
<td>(0.455)</td>
<td>(0.455)</td>
<td>(0.468)</td>
<td>(0.518)</td>
</tr>
<tr>
<td>log(Adv_{t-1}/Sales_{t-1})</td>
<td>0.814***</td>
<td>0.815***</td>
<td>0.802***</td>
<td>0.802***</td>
</tr>
<tr>
<td></td>
<td>(0.00732)</td>
<td>(0.00732)</td>
<td>(0.00789)</td>
<td>(0.00789)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.943***</td>
<td>-0.942***</td>
<td>-0.918***</td>
<td>-1.469***</td>
</tr>
<tr>
<td></td>
<td>(0.141)</td>
<td>(0.141)</td>
<td>(0.132)</td>
<td>(0.352)</td>
</tr>
<tr>
<td>Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Industry FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>State FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>24082</td>
<td>24082</td>
<td>22065</td>
<td>22065</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.77</td>
<td>0.77</td>
<td>0.76</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Notes: Data from Compustat from 1950 to 2009, Wilson (2009) and Falato and Sim (2014). Firm age is measured as the elapsed time since the first observation in the data. Our measure of financial constraints is sale minus purchases of common and preferred stock, divided by sales. Standard errors are clustered at the firm level (in parentheses). Significance level: * 10%; ** 5%; *** 1%.
7.1 Advertising Tax

We introduce a tax on advertising expenditure along the BGP. Using Equation (14), the firm’s optimal advertising choice maximizes:

\[
\max_{\{m_j\}} \sum_j \left[ \tilde{\pi} \cdot (q_j + q_j d_j) - (1 + \xi) m_j \right] \quad \text{s.t.} \quad q_j d_j = \theta_j m_j \hat{\xi} n^n
\]

where \( \xi \geq 0 \) is the tax rate, which we assume to be time- and size-invariant. Namely, for every unit of expenditure in advertising, the firm must pay an additional cost of \( \xi \). The tax is observationally equivalent to a decrease in \( \theta_j \),\(^{57}\) and its proceeds are rebated lump-sum to households.

Figure 12 plots, for different levels of the tax, the percentage change (with respect to \( \xi = 0 \)) in the growth, entry, internal and external innovation rates (the latter for firms of different size \( n \)) of the calibrated economy. We observe that the advertising tax is equivalent to a decrease in advertising efficiency, which increases the growth rate of the economy (as seen in Figure 8). However, the tax is considerably ineffective in the sense that the change in growth is small.

\(^{57}\) In particular, \( \theta_{\text{post-tax}} = \theta_{\text{pre-tax}} (1 + \xi)^{-\xi} \).
There are two main explanations for this result. First, the elasticity of sales to advertising expenditure is low, which implies that firms’ response to the change in the advertising cost is timid. In particular, they only very modestly substitute advertising spending for R&D spending, and overall the economy’s growth rate responds very little. Secondly, the advertising tax deters entry and shifts mass toward large firms in the firm distribution, both of which are a force toward less growth. This is because when advertising becomes less efficient, the dynamic incentives for entrants to acquire an initial good on which to start obtaining spillover effects from advertising in the future are smaller. This channel mutes the positive growth effects of the tax.

Overall, we find that even though the advertising sector is growth and welfare detrimental, directed corrective taxation has limited effective power. While in the previous section we found that banning advertising (in the sense of having $\theta = 0$) would have a significant impact on the growth rate of the economy (namely, a 0.64 percentage-point increase), here we find that implementing such a policy would require a “virtually unfeasible” tax rate. In fact, even to obtain small though non-negligible effects on growth, tax rates would need to be unreasonably high. For instance, a 0.1 percentage-point increase in $g$ would require a tax rate around 910%.

7.2 R&D Subsidies

Next, we analyze the effect of R&D subsidies with and without the presence of advertising. A study of this type of industrial policy in the context of an endogenous growth model of firm dynamics is not new to our work. Akcigit (2009) and Acemoglu, Akcigit, Bloom, and Kerr (2013), among others, find that R&D subsidies may increase growth, particularly when targeted toward entrants and small incumbent innovative-intensive firms. This policy scheme is typically welfare-improving because it has the potential of correcting for inefficiencies stemming from R&D: namely, in the decentralized equilibrium, firms fail to appropriate the full social value of their innovations, as they do not internalize that their current innovations will be embedded within the goods that will be passed on to other firms in the future through creative destruction.\footnote{The inefficiency in Schumpeterian models arising from a business stealing effect is usually shown to be dominated by positive externalities. See for instance Jones and Williams (2000), Alvarez-Pelaez and Groth (2005) and Bloom, Schankerman, and Van Reenen (2013) for evidence that the private return to R&D is lower than its social return.}
In our calibrated economy, advertising poses a second source of distortions, as it impedes the economy from reaching its full growth potential by altering innovation decisions. Here, we intend to explore the implications of a subsidy on R&D, when the advertising channel provides the necessary spillover effect to explain the deviation from Gibrat’s law seen in the data. We then compare this baseline economy to another one in which there are decreasing returns to scale in R&D and no advertising.

**Calibration Without Advertising and DRTS in R&D**  We shall compare two subsidized economies. The first economy corresponds to our baseline calibration of Section 5.1, in which the spillover effect on advertising is active and there exist constant returns to scale in R&D. The second economy corresponds to a calibration of the model in which advertising is shut down (e.g. $\theta = 0$) and returns to R&D need not scale proportionally with firm size (i.e., we do not impose that $\psi + \sigma = 1$). In this new specification, we have 2 fewer parameters to estimate ($\zeta$, $\theta$ and $\eta$ are now not calibrated, but $\sigma$ is), and 5 as opposed to 7 moments to target (as we lose the advertising moments).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>0.02</td>
<td>Time discount rate</td>
<td>Standard</td>
</tr>
<tr>
<td>$\psi$</td>
<td>2</td>
<td>Elasticity of R&amp;D</td>
<td>Akcigit and Kerr (2015)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1645</td>
<td>Profitability ratio</td>
<td>Compustat</td>
</tr>
</tbody>
</table>

Internally identified

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>0.0226</td>
<td>Innovation step</td>
<td>.</td>
</tr>
<tr>
<td>$\hat{\chi}$</td>
<td>0.0116</td>
<td>Internal R&amp;D scale param.</td>
<td>.</td>
</tr>
<tr>
<td>$\tilde{\chi}$</td>
<td>3.2280</td>
<td>External R&amp;D scale param.</td>
<td>.</td>
</tr>
<tr>
<td>$\nu$</td>
<td>1.2351</td>
<td>Entrant’s R&amp;D scale param.</td>
<td>.</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>-0.9411</td>
<td>Degree of DRTS in R&amp;D</td>
<td>.</td>
</tr>
</tbody>
</table>

| Implied $\psi + \sigma$ | 1.0589 |

Table 10: Full set of calibrated parameters in the model without advertising and with decreasing returns in R&D.

Table 10 presents the parameter values for the calibrated model without advertising, and Table 11 presents its performance in terms of moment matching (next to the results from the baseline calibration with advertising that we presented before in Table 6). As expected, absent
the advertising channel, we need decreasing returns to scale in R&D (i.e., \( \psi + \sigma > 1 \)) in order to match the deviations from constant growth and R&D intensity. In particular, we find \( \sigma = -0.9411 < 1 \) (so that \( \tilde{\psi} + \sigma = 1.0589 \)) which implies that a 10-product firm would find it 14.5% more expensive to grow by a given rate than it would be for 10 combined firms of one product each. In terms of the micro-level empirical facts, we now have a hard time matching the deviation from Gibrat’s law. The baseline calibration of Section 5.1 with advertising spillover effect and constant returns in R&D was not only able to closely match both the deviation from Gibrat’s law and from constant R&D intensity, but also the advertising regression coefficient.

By analyzing the different effects of an R&D subsidy on each of these two economies, we intend to illustrate the differences in terms of growth impact when our advertising spillover effect is active. Ultimately, we intend to determine if the dynamic interaction between advertising and R&D alters the effectiveness of these policies, and generally we seek to shed light on whether the source of decreasing returns to scale is relevant for the effectiveness of economic policy.

<table>
<thead>
<tr>
<th>Moment</th>
<th>Model with ADV</th>
<th>Model w/o ADV</th>
<th>Data</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate moments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average growth rate</td>
<td>( g )</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Firm entry rate</td>
<td>( x_e / F )</td>
<td>0.101</td>
<td>0.097</td>
<td>0.098</td>
</tr>
<tr>
<td>Average R&amp;D/Sales</td>
<td>( \sum_n F \mu_n R_n / (py) )</td>
<td>0.153</td>
<td>0.097</td>
<td>0.102</td>
</tr>
<tr>
<td>Average R&amp;D/ADV</td>
<td>( \sum_n F \mu_n R_n / M_n )</td>
<td>24.15</td>
<td>.</td>
<td>26.40</td>
</tr>
</tbody>
</table>

Regression coefficients

<table>
<thead>
<tr>
<th></th>
<th>( \beta_{\Delta \text{sales/sales}} )</th>
<th>( \beta_{\text{rd/sales}} )</th>
<th>( \beta_{\text{adv/sales}} )</th>
<th>( \beta_{\text{rd/adv}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gibrat’s coefficient</td>
<td>-0.0326</td>
<td>-0.0217</td>
<td>-0.0325</td>
<td>Table 1</td>
</tr>
<tr>
<td>R&amp;D intensity coeff.</td>
<td>-0.1030</td>
<td>-0.1041</td>
<td>-0.1035</td>
<td>Table 2</td>
</tr>
<tr>
<td>ADV intensity coeff.</td>
<td>-0.0353</td>
<td>.</td>
<td>-0.0317</td>
<td>Table 3</td>
</tr>
<tr>
<td>R&amp;D/ADV coefficient</td>
<td>-0.0677</td>
<td>.</td>
<td>-0.0719</td>
<td>Table 4</td>
</tr>
</tbody>
</table>

Table 11: Targeted moments: Model #1: advertising and constant returns to scale in R&D; Model #2: No advertising and decreasing returns to scale in R&D. No advertising and decreasing returns to scale in R&D.

Notes: Sales is used as the measure of firm size. \( R_n \) stands for the total R&D expenditures of a firm of size \( n \), defined in Equation (3). Similarly, \( M_n \) is total advertising expenditures at the firm level, defined in Equation (16).

Results  We assume that a time- and size-invariant subsidy \( s \in (0, 1) \) is given to each type of firm conducting R&D (including both entrants and incumbents). This subsidy is financed by a lump-sum tax on households. The R&D cost functions now read:
\[ R_e(x_e) = (1 - s)\nu \tilde{Q}_x x_e; \quad R_z(z_j) = (1 - s)\tilde{\chi}_j \hat{\psi}_j; \quad R_x(X_n, n) = (1 - s)\tilde{\chi} Q X_n^\psi n^\sigma \]

Figures 13 and 14 show the results. First, Figure 13 shows the percentage change, with respect to the no-subsidy allocation \((s = 0)\), in growth, entry and internal innovation rates, under both calibrated economies and for different levels of the subsidy. In line with previous studies, we find that the subsidy increases growth. This is because innovation becomes cheaper, which fosters entry and raises R&D expenditure for all firm sizes. In fact, the percentage change in the growth rate is increasingly higher the more generous the subsidies are. More importantly, as seen in Table 12, we find that the subsidy has a bigger impact on growth in the economy with advertising and constant returns to scale in the R&D technology (red dashed line in Figure 13). For instance, in the economy with advertising, a 50\% subsidy increases the growth rate to \(g = 2.79\%\), while in the economy without advertising a subsidy of the same size increases the rate to \(g = 2.66\%\), i.e. 13 percentage points less. This differential gets bigger with the subsidy. For instance, a subsidy of 75\% increases the growth rate by 29 percentage points more in the economy with advertising.

<table>
<thead>
<tr>
<th>Subs.</th>
<th>Calibration with ADV and CRTS</th>
<th>Calibration with DRTS and no ADV</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>25%</td>
<td>2.31%</td>
<td>2.25%</td>
</tr>
<tr>
<td>50%</td>
<td>2.79%</td>
<td>2.66%</td>
</tr>
<tr>
<td>75%</td>
<td>3.85%</td>
<td>3.56%</td>
</tr>
</tbody>
</table>

Table 12: Growth rates under both calibrated economies, for different subsidies. Left column: baseline calibration with advertising and constant returns to scale in R&D. Right column: calibration with decreasing returns to scale in R&D and no advertising \((\theta = 0)\).

In the economy with advertising, the subsidy has a much larger positive impact on entry, as R&D is cheaper and smaller firms benefit more from innovating through the advertising spillover channel. The impact of the subsidy on the internal innovation rate is also stronger. Both of these contribute to growth disproportionately more than they do in the economy without advertising.

Another noteworthy effect lies in Figure 14, which plots the change in the external innovation rate for different firm sizes for a change from \(s = 0\%\) to \(s = 20\%\). When there are constant
returns to scale in R&D and the spillover effect in advertising is active (red dashed line), firms are more responsive to the subsidy in terms of innovation investment than if there were no advertising and the R&D technology had decreasing returns to scale (blue solid line). This is because in the economy with advertising, all firms are equally efficient in producing innovations. Instead, in the economy without advertising, larger firms are less efficient in conducting R&D, which dampens their response to the subsidy. This is particularly true for larger firms as the difference in external R&D response to a given change in subsidies gets larger with firm size.
Consequently, absent advertising, subsidies should be larger in magnitude in order to reach the effectiveness that would be attained in an economy in which deviations from constant growth are not uniquely explained by technological differences in innovation capacities. More generally, our results suggest that identifying the source of technical efficiency in producing innovation across firm sizes has very relevant implications for the impact of industrial policies directed to increase growth.

8 Conclusion

In this paper, we have presented a model of firm dynamics and endogenous growth through product innovation that explicitly incorporates advertising decisions by firms. In modeling advertising, we have been inspired by observations from the empirical marketing literature, specifically the existence of advertising spillovers across goods. In our calibrated model, smaller firms benefit more in terms of decreased advertising cost from brand extension. Such extensions are achieved through external R&D, and smaller firms are, therefore, relatively more concerned with conducting innovative activities, as these firms benefit more from advertising additional product lines due to the spillover effect. This mechanism generates the empirical observation that R&D intensity is decreasing with firm size even in the absence of differences in R&D technology across firm size.

We have shown that the model can qualitatively as well as quantitatively match empirical regularities across sizes in sales growth, R&D intensity and advertising intensity. On the basis of our calibration, we have shown that advertising is detrimental to growth and welfare because it crowds out innovation investment within the firm. Even though advertising increases aggregate consumption contemporaneously by shifting preferences toward certain goods, it substantially decreases long term growth as it also shifts incentives away from productivity-enhancing activities. Key to this argument is that R&D and advertising are strategic substitutes in the equilibrium of the calibrated model. We confirm this substitution empirically by exploiting exogenous cross-state and cross-time variation in the cost of innovation arising from changes in the tax treatment of R&D in the U.S. from the 1980s onwards.
Based on these results, we have proposed a set of growth-enhancing industrial policies. First, we have shown that advertising taxes are relatively ineffective because the elasticity of advertising to sales is low and outside firms are deterred from entry. Both of these mute the effects of the tax on growth. Second, R&D subsidies have a higher impact on growth in an economy with advertising and constant returns to scale in R&D than in one in which there is no advertising and where larger firms are technologically less efficient in conducting innovative activities. Identifying the source of decreasing R&D intensity with firm size is then an important question and remains an interesting avenue for future research.
References


A  Data

A.1  Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>ADV and R&amp;D firms</th>
<th>All firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>St. dev.</td>
</tr>
<tr>
<td>Sales (million U.S.D.)</td>
<td>2539.6</td>
<td>12222.5</td>
</tr>
<tr>
<td>Age (in years)</td>
<td>14.6</td>
<td>12.72</td>
</tr>
<tr>
<td>Equity Fin./Sales</td>
<td>0.269</td>
<td>6.09</td>
</tr>
<tr>
<td>R&amp;D (million U.S.D.)</td>
<td>113.5</td>
<td>580.8</td>
</tr>
<tr>
<td>ADV (million U.S.D.)</td>
<td>73.2</td>
<td>402.9</td>
</tr>
<tr>
<td>R&amp;D/Sales</td>
<td>0.102</td>
<td>0.136</td>
</tr>
<tr>
<td>ADV/Sales</td>
<td>0.037</td>
<td>0.068</td>
</tr>
<tr>
<td>R&amp;D/ADV</td>
<td>26.4</td>
<td>229.56</td>
</tr>
</tbody>
</table>

| Share of firms doing R&D| 31.60% |
| Share of sales from firms doing R&D | 44.20% |
| Share of firms doing ADV  | 33.90% |
| Share of sales from firms doing ADV | 36.20% |

Table A.1: Descriptive statistics. Compustat data from 1980 to 2015. Means and standard deviations in our sample of firms performing R&D and advertising vs. all firms in Compustat. Reported shares are computed with respect to the whole Compustat universe.

A.2  Additional Tables
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sales</td>
<td>Sales</td>
<td>Sales</td>
<td>EBIT</td>
<td>EBIT</td>
<td>EBIT</td>
<td>OIBD</td>
<td>OIBD</td>
<td>OIBD</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>8.998***</td>
<td>8.976***</td>
<td>8.094***</td>
<td>1.151***</td>
<td>1.134***</td>
<td>1.200***</td>
<td>1.757***</td>
<td>1.746***</td>
<td>1.736***</td>
</tr>
<tr>
<td></td>
<td>(1.526)</td>
<td>(1.532)</td>
<td>(1.359)</td>
<td>(0.264)</td>
<td>(0.267)</td>
<td>(0.294)</td>
<td>(0.242)</td>
<td>(0.243)</td>
<td>(0.265)</td>
</tr>
<tr>
<td>Advertising</td>
<td>6.788***</td>
<td>6.770***</td>
<td>7.933***</td>
<td>0.849***</td>
<td>0.835***</td>
<td>0.687***</td>
<td>1.228***</td>
<td>1.218***</td>
<td>1.214***</td>
</tr>
<tr>
<td></td>
<td>(1.713)</td>
<td>(1.709)</td>
<td>(1.976)</td>
<td>(0.265)</td>
<td>(0.260)</td>
<td>(0.266)</td>
<td>(0.336)</td>
<td>(0.333)</td>
<td>(0.363)</td>
</tr>
<tr>
<td></td>
<td>(18.03)</td>
<td>(18.00)</td>
<td>(17.53)</td>
<td>(2.160)</td>
<td>(2.118)</td>
<td>(2.070)</td>
<td>(3.086)</td>
<td>(3.076)</td>
<td>(2.914)</td>
</tr>
<tr>
<td></td>
<td>(17.98)</td>
<td>(16.61)</td>
<td>(2.039)</td>
<td>(2.106)</td>
<td>(2.637)</td>
<td>(2.655)</td>
<td>(2.637)</td>
<td>(2.655)</td>
<td>(3.225)</td>
</tr>
<tr>
<td></td>
<td>(7.514)</td>
<td>(1.550)</td>
<td>(1.550)</td>
<td>(1.550)</td>
<td>(1.550)</td>
<td>(1.550)</td>
<td>(1.550)</td>
<td>(1.550)</td>
<td>(1.550)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1417.4**</td>
<td>-1447.9**</td>
<td>-1351.5**</td>
<td>64.58</td>
<td>41.63</td>
<td>56.56</td>
<td>-29.25</td>
<td>-45.39</td>
<td>-15.04</td>
</tr>
<tr>
<td></td>
<td>(555.6)</td>
<td>(564.7)</td>
<td>(533.3)</td>
<td>(127.2)</td>
<td>(131.0)</td>
<td>(139.3)</td>
<td>(99.33)</td>
<td>(102.3)</td>
<td>(115.3)</td>
</tr>
<tr>
<td>Year dummies</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Industry FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>25619</td>
<td>25619</td>
<td>23427</td>
<td>25617</td>
<td>25617</td>
<td>23425</td>
<td>25604</td>
<td>25604</td>
<td>23414</td>
</tr>
<tr>
<td>R²</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.58</td>
<td>0.58</td>
<td>0.56</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
</tr>
</tbody>
</table>

**Table A.2:** Sale and profit regressions.

*Notes:* “EBIT” stands for Earnings Before Interest and Tax. “OIBD” stands for Operating Income Before Depreciation. Sales, EBIT, OIBD, R&D and advertising expenditures are in thousands of U.S.D. See Table 1 for details on the sample selection. Standard errors are clustered by firm (in parentheses). Legend: * 10%; ** 5%; *** 1%.
Table A.3: Facts 1 to 4 with assets and employment as measures of size.

Notes: Sales, R&D and advertising expenditures are in thousands of U.S. D. Employment is in thousands of employees. See Table 1 for details on the sample selection. Standard errors are clustered by firm (in parentheses). Legend: * 10%; ** 5%; *** 1%.
B Derivations and Proofs

B.1 Value Functions

The value function for a typical incumbent firm with intrinsic quality portfolio \( q \) and \( n \) products at time \( t \), denoted by \( V_{n,t}(q) \), is given by:

\[
V_{n,t}(q) = \max_{\{z_j,t, m_j,t\}} \left\{ \sum_{q_j \in q} \left( \bar{\pi}_t(q_j + \phi_j) - \bar{\chi} z^\sigma_{j,t} q_j - m_{j,t} \right) \Delta t - \bar{\chi} n^\sigma + \bar{\psi} x_t \bar{Q}_t \Delta t \right. \\
+ e^{-r_t + \Delta t} \left( \sum_{q_j \in q} \left[ z_j,t \Delta t + o(\Delta t) \right] V_{n,t+\Delta t}(q \setminus \{q_j\} \cup \{q_j(1 + \lambda^I)\}) \right. \\
+ \left. \sum_{q_j \in q} \left[ \tau_t \Delta t + o(\Delta t) \right] V_{n-1,t+\Delta t}(q \setminus \{q_j\}) \right. \\
+ \left[ n x_t \Delta t + o(\Delta t) \right] E_j \left\{ V_{n+1,t+\Delta t}(q \cup \{q_j(1 + \lambda^E)\}) \right\} \\
+ \left. \left[ 1 - \sum_{q_j \in q} z_j,t \Delta t - \sum_{q_j \in q} \tau_t \Delta t - n x_t \Delta t - o(\Delta t) \right] V_{n,t+\Delta t}(q) \right\} + o(\Delta t)
\]

where \( o(\Delta t) \) has the property \( \lim_{\Delta \to 0} \frac{o(\Delta t)}{\Delta t} = 0 \). Here, we use the fact that, for each Poisson arrival rate \( k \in \{z, x, \tau\} \), the term \( k \Delta t + o(\Delta t) \) (respectively, \( 1 - k \Delta t - o(\Delta t) \)) approximates the probability of exactly one Poisson event (respectively, zero Poisson events) within an interval of short length (i.e., for small \( \Delta > 0 \)). The probability of two or more events is equal to \( o(\Delta t) \) in the limit as \( \Delta \to 0 \).

The interpretation of the right-hand side of the equation is as follows. The first line includes flow profits for each good in the firm’s portfolio, net of R&D and advertising costs; the second, third, fourth and fifth lines include the different scenarios that can arise at the next instant, all of which are exponentially discounted by the interest rate \( r_{t+\Delta t} \Delta t \): an internal innovation on some good \( j \), with overall probability \( \sum_{q_j \in q} \left[ z_j,t \Delta t + o(\Delta t) \right] \), which advances the quality portfolio by a factor of \( \lambda^I \) on the newly innovated product (second line); external innovations, either by an external firm on some currently held good, which occurs with overall probability \( \sum_{q_j \in q} \left[ \tau_t \Delta t + o(\Delta t) \right] \) (third line), or by the firm in question over a randomly drawn good from the \([0, 1]\) continuum, which occurs with overall probability \( \sum_{q_j \in q} \left[ n x_t \Delta t + o(\Delta t) \right] \) (fourth line).
With the remaining probably, none of these events occurs, and both the quality portfolio and the number of products of the firm stay the same (fifth line).

To obtain Equation (22), simply subtract \( e^{-r_t + \Delta t} V_{n,t}(q) \) from both sides, divide every term by \( \Delta t \), and take the continuous-time limit as \( \Delta \to 0 \) (using the fact that \( \lim_{\Delta \to 0} o(\Delta) \Delta = 0 \) and \( \lim_{\Delta \to 0} \frac{1 - e^{-r_t + \Delta t} \Delta}{\Delta} = r_t \)).

As for entrants, the value function is

\[
V_{0,t} = \max_{x_{e,t} > 0} \left\{ -\nu x_e \bar{Q} \Delta t + e^{-r_t + \Delta t} \left( x_{e,t} \Delta t + o(\Delta t) \right) \mathbb{E}_j \left\{ V_{1,t+\Delta t}(\{q_j(1 + \lambda^E)\}) \right\} 
+ \left[ 1 - x_{e,t} \Delta t - o(\Delta t) \right] V_{0,t+\Delta t} \right\} + o(\Delta t)
\]

where \( V_{0,t} \) denotes firm value at \((n, q) = (0, \emptyset)\). The first line includes the cost of entry (first term), and the value of the firm in the event of entry at the next instant, which occurs with probability \( x_{e,t} \Delta t + o(\Delta t) \). The second line shows that the firm remains with zero products with the complementary probability. Once again, the probability of two or more such events occurring simultaneously is a higher-order term inside \( o(\Delta t) \).

To obtain Equation (23), once again subtract \( e^{-r_t + \Delta t} V_{0,t} \) from both sides, divide through by \( \Delta t \), and take the limit as \( \Delta \to 0 \).

### B.2 Proof of Proposition 1

Assume that \( z_{j,t} = z_t, \forall j \in [0, 1] \), a result that we prove independently in Proposition 3. At any time \( t \), aggregate quality is given by:

\[
\tilde{Q}_{t+\Delta t} = \left[ \tau_t \Delta t + o(\Delta t) \right] (1 + \lambda^E) \tilde{Q}_t + \left[ z_t \Delta t + o(\Delta t) \right] (1 + \lambda^I) \tilde{Q}_t + \left[ 1 - \tau_t \Delta t - z_t \Delta t - o(\Delta t) \right] \tilde{Q}_t + o(\Delta t)
\]

The interpretation of this equation is straightforward: after an instant of length \( \Delta t \), aggregate intrinsic quality either increases by a fixed factor, due to an external innovation or an internal
one (with probability $\tau t\Delta t + o(\Delta t)$ and $z t\Delta t + o(\Delta t)$, respectively), or it remains the same (with the complementary probability). The probability of two or more innovations is a second-order event. Subtracting $\dot{Q}_t$ from both sides, dividing through by $\Delta t$, taking the limit as $\Delta \to 0$ and using that $\lim_{\Delta \to 0} \frac{o(\Delta t)}{\Delta t} = 0$, gives $\dot{Q}_t = \tau \lambda^E \dot{Q}_t + z \lambda^I \dot{Q}_t$. Therefore,

$$g = \tau \lambda^E + z \lambda^I$$

as we sought to show. □

### B.3 Proof of Proposition 2

Let $\mu_n$ denote the equilibrium share of incumbent firms that own $n \geq 1$ product lines, such that $\mu_n \in [0, 1]$, $\forall n$, and $\sum_{n=1}^{+\infty} \mu_n = 1$. The invariant distribution must satisfy the following flow equations:

<table>
<thead>
<tr>
<th># Products</th>
<th>Inflows</th>
<th>Outflows</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 0$</td>
<td>$F \mu_1 \tau$</td>
<td>$= x_e$</td>
</tr>
<tr>
<td>$n = 1$</td>
<td>$F \mu_2 2\tau + x_e$</td>
<td>$= F \mu_1 (x_1 + \tau)$</td>
</tr>
<tr>
<td>$n \geq 2$</td>
<td>$F \mu_{n+1} (n + 1) \tau + F \mu_{n-1} (n - 1) x_{n-1}$</td>
<td>$= F \mu_n n (x_n + \tau)$</td>
</tr>
</tbody>
</table>

The left-hand (or right-hand) side of these equalities describes the mass of firms that enters into (or exits out of) the state identified by the first column. For instance, for $n \geq 2$, inflows are given by the mass of size-$(n + 1)$ incumbent firms that lost a product through creative destruction, and the mass of size-$(n - 1)$ incumbent firms that gained one through external innovation. Similarly, outflows are given by the mass of size-$n$ incumbents that either lost a product to another incumbent, or gained one thanks to external innovation. Because innovation allows firms to gain or lose only one product at a time, potential entrants (entering at rate $x_e$) only feature as a flow out of $n = 0$, or into $n = 1$.

We prove that Equation (26) is the solution to the above system of flow equations by use
of mathematical induction. Specializing the formula to \( n = 1 \), and using the convention that \( \prod_{i=1}^{0} x_i = 1 \), we have

\[
\mu_1 = \frac{x_e}{F\tau}
\]

which is true by the first flow equation. For \( n = 2 \), Equation (26) gives

\[
\mu_2 = \frac{x_e x_1}{2F\tau^2}
\]

which can be written as \( 2F\tau^2 \mu_2 = x_1 x_e \). Adding \( x_e\tau \) to both sides and dividing through by \( F\tau \) we obtain \( \mu_2 2\tau + \frac{x_e}{F} = \frac{x_e}{F} \left( \frac{x_1}{\tau} + 1 \right) \). Rearranging, we obtain exactly the flow equation for \( n = 2 \).

Having shown that Equation (26) holds for \( n = 1 \) and \( n = 2 \), it is enough to show that it holds for \( n + 1 \), assuming it does for \( n \) and \( n - 1 \), where \( n \geq 2 \) is arbitrary. Accordingly, suppose \( \mu_{n-1} = \frac{x_e}{F} \frac{\prod_{i=1}^{n-2} x_i}{(n-1)\tau^{n-1}} \) and \( \mu_n = \frac{x_e}{F} \frac{\prod_{i=1}^{n-1} x_i}{n\tau^n} \). Plugging these into the third flow equation, we have

\[
F\mu_{n+1}(n+1)\tau + x_e \frac{\prod_{i=1}^{n-2} x_i}{\tau^{n-1}} x_{n-1} = x_e \frac{\prod_{i=1}^{n-1} x_i}{\tau^n} (x_n + \tau)
\]

Simplifying,

\[
F\mu_{n+1}(n+1)\tau = x_e \left( \frac{\prod_{i=1}^{n-1} x_i}{\tau^n} (x_n + \tau) - \frac{\prod_{i=1}^{n-2} x_i}{\tau^{n-1}} x_{n-1} \right)
\]

\[
= x_e \left( \frac{\prod_{i=1}^{n} x_i}{\tau^n} + \tau \frac{\prod_{i=1}^{n-1} x_i}{\tau^n} - \frac{\prod_{i=1}^{n-1} x_i}{\tau^{n-1}} \right)
\]

\[
= x_e \frac{\prod_{i=1}^{n} x_i}{\tau^n}
\]

which implies that \( \mu_{n+1} = \frac{x_e}{F} \frac{\prod_{i=1}^{n} x_i}{(n+1)\tau^{n+1}} \), what we wanted to show. ∎

### B.4 Proof of Proposition 3

We find \( \Gamma \) and \( \{\Upsilon_n\}_{n=1}^{\infty} \) using the method of undetermined coefficients. Plugging the guess

\[
V_n(q_f) = \Gamma \sum_{q_j \in q_f} q_j + \Upsilon_n \bar{Q}
\]

into (22), we get that
\begin{align*}
r\Gamma \sum_{q_j \in \mathcal{q}_f} q_j + r\Upsilon_n \hat{Q} &= \max_{x \in [0,x]} \left\{ \sum_{q_j \in \mathcal{q}_f} \left[ z_j \lambda' q_j + \tilde{\pi} q_j + \tau \left( (\Upsilon_{n-1} - \Upsilon_n) \hat{Q} - \Gamma q_j \right) - \tilde{\chi} z_j \hat{q}_j \right] \right. \\
&\quad + nx \left( \Gamma \hat{Q}(1 + \lambda E) + (\Upsilon_{n+1} - \Upsilon_n) \hat{Q} \right) - \tilde{\chi} n^{\sigma+\psi} x^{\tilde{\psi}} \hat{Q} + \gamma Q n^{\frac{n}{-\zeta}} \right\} + \Upsilon_n \hat{Q} g
\end{align*}

Equating the terms with \( q_j \) and \( \hat{Q} \) gives the following pair of optimization problems:

\begin{align*}
(q_j) : \quad r\Gamma &= \max_{z_j} \left\{ \tilde{\pi} + z_j \lambda' - \tau \Gamma - \tilde{\chi} z_j \right\} \\
(\hat{Q}) : \quad (r - g) \Upsilon_n &= \max_{x_n} \left\{ (\Upsilon_{n-1} - \Upsilon_n) n \tau + nx_n \left( \Gamma(1 + \lambda E) + \Upsilon_{n+1} - \Upsilon_n \right) - \tilde{\chi} n^{\sigma+\psi} x_n^{\tilde{\psi}} + \gamma n^{\frac{n}{-\zeta}} \right\}
\end{align*}

The first-order conditions are:

\begin{align*}
z_j &= \left( \frac{\Gamma \lambda'}{\tilde{\psi} \hat{\chi}} \right)^{\frac{1}{\psi - 1}} \\
x_n &= n^{\frac{1 - \psi - \tilde{\psi}}{\psi - 1}} \left( \frac{\Gamma(1 + \lambda E) + \Upsilon_{n+1} - \Upsilon_n}{\tilde{\psi} \hat{\chi}} \right)^{\frac{1}{\psi - 1}}
\end{align*}

respectively. Assuming that there is positive entry in equilibrium (\( x_e > 0 \)), we can exploit the free-entry condition \( V_0 = 0 \) in (23) to get that

\begin{align*}
\Gamma &= \frac{\nu - \Upsilon_1}{1 + \lambda E}
\end{align*}

This means that the optimal internal R&D investment by incumbents is

\begin{align*}
z_j &= \left( \frac{\lambda'(\nu - \Upsilon_1)}{\tilde{\psi} \hat{\chi}(1 + \lambda E)} \right)^{\frac{1}{\psi - 1}}
\end{align*}

so \( z_j = z, \forall j \in [0, 1] \). Back into the optimality condition for \( z \), we can obtain the implied rate of creative destruction:
\[ \tau = \frac{1}{\nu - \Upsilon_1} \left[ \tilde{\pi} - \hat{\chi} \left( \frac{\lambda^I (\nu - \Upsilon_1)}{\hat{\chi} \hat{\psi} (1 + \lambda^E)} \right)^{\hat{\psi}^{-1}} \right] - \frac{\rho}{1 + \lambda^E} \]

where we have used that \( g = r - \rho \) from the Euler equation, and \( g = \tau \lambda^E + z \lambda^I \). It remains to find an expression for \( \Upsilon_n \). Using free-entry, we know

\[ x_n = n^{1-\sigma-\hat{\psi}^{-1}} \left( \frac{\nu - \Upsilon_1 + \Upsilon_{n+1} - \Upsilon_n}{\tilde{\psi} \hat{\chi}} \right)^{\hat{\psi}^{-1}} \]

Back into the second maximization problem, we get the second-order difference equation

\[ (r - g) \Upsilon_n = (\Upsilon_{n-1} - \Upsilon_n) n \tau + \gamma n^{n/(1-\zeta)} + \tilde{\chi}(\tilde{\psi} - 1) n^{\frac{-\sigma}{\psi^{-1}}} \left[ \frac{\nu - \Upsilon_1 + \Upsilon_{n+1} - \Upsilon_n}{\tilde{\psi} \hat{\chi}} \right]^{\hat{\psi}^{-1}} \]

Solving for \( \Upsilon_{n+1} \) gives (28). Using \( V_0 = 0 \) and the guess \( V_n = \Gamma \sum_j q_j + \Upsilon_n \bar{Q} \), it is clear that the boundary condition for this difference equation must then be \( \Upsilon_0 = 0 \). □

C Model Extensions

C.1 Patents and Major Innovations

In our baseline model, smaller firms are more innovative because of the advertising spillover effect, in spite of there being constant returns to scale in R&D. Besides differences in R&D intensity across size, however, the literature has emphasized other motives why smaller firms might be relatively more efficient in conducting innovations. One such literature has focused on patent behavior. Firstly, entrants and smaller firms typically produce relatively more major and radical innovations (when the quality of a patent is based on the number of external citations that it receives).\(^{59}\) Secondly, these firms also tend to patent relatively more on average. In this section, we show that a simple extension of our baseline model with advertising can deliver both of these additional facts even in the absence of decreasing returns to scale in R&D.

\(^{59}\)Indeed, Akcigit and Kerr (2015) show that the fraction of a firm’s patent in the top patent quality decile is decreasing with firm size.
We first demonstrate how to obtain the first stylized fact, namely that smaller firms and entrants generate more major technological improvements. Intuitively, since in our specification these firms invest relatively more than larger firms in external innovation because they benefit marginally more from the advertising effect, the result directly obtains if the innovation step size for external innovations is larger than that of internal innovations.

Let us assume that each innovation creates a new patent that potentially cites other patents that exist at the time the new patent is introduced. Following Akcigit and Kerr (2015), each innovation belongs to a technological cluster, and there exist two types of technological advances: follow-up and major advances. A major advance in a production line creates a whole new cluster of innovation, while a follow-up innovation belongs to the same cluster as the patent that it improves upon. Each patent within a technological cluster is assumed to cite all the previous innovations in the same cluster with some positive probability. Consequently, major advances in the model do not cite any other existing patent and potentially receive citations from follow-up innovations within the same cluster. Once a new major innovation creates a new technological cluster, the cluster that it replaces receives no more citations. Likewise, follow-up innovations can also receive citations from subsequent follow-up innovations in the same cluster. However, on average, major advances receive more citations than follow-up innovations.

Formally, we extend the baseline model by allowing external innovations to result in major technological advances with some probability. Whereas internal R&D can only result in a follow-up innovation, with step size $\lambda^I > 0$, external innovation can either be a follow-up within an existing technological cluster, or be a major advance and create a new technological cluster altogether. Let $\omega$ denote the probability with which a successful external innovation leads to a major technological advance. In this case, the step size for quality improvement is equal to $\lambda^H > \lambda^I$. With the remaining probability $(1 - \omega)$, the successful external innovation is a follow-up, which leads to a step size $\lambda^L \in (0, \lambda^H]$. In sum, for any product line $j$ and a small interval $\Delta t > 0$, intrinsic quality is given by
\[
q_{j,t+\Delta t} = q_{jt} + \begin{cases} 
\lambda^H q_{jt} & \text{w.prob. } \omega \tau_t \Delta t + o(\Delta t) \quad \text{[External, major advance]} \\
\lambda^L q_{jt} & \text{w.prob. } (1-\omega) \tau_t \Delta t + o(\Delta t) \quad \text{[External, follow-up]} \\
\lambda^I q_{jt} & \text{w.prob. } z_{jt} \Delta t + o(\Delta t) \quad \text{[Internal]} \\
0 & \text{w.prob. } 1 - \tau_t \Delta t - z_{jt} \Delta t - o(\Delta t) \quad \text{[No innovations]} 
\end{cases}
\]

Therefore, the average quality improvement from a successful external innovation is equal to \( \lambda^E \equiv \omega \lambda^H + (1-\omega) \lambda^L \), which means that the growth and creative destruction rates of the economy are still given by (24) and (25), respectively. Therefore, our baseline model can be thought of a special case of the extended model with patents, and none of the results derived in the baseline case will change. Additionally, the model now predicts that smaller firms, while exhibiting a higher R&D intensity because of the advertising spillover effect, also produce relatively more major innovations.

Let us now show how the model can deliver the prediction that smaller firms tend to patent relatively more on average, and that their patents are of higher quality (when quality is measured by the number of external citation that it receives). For this, we can characterize expected patent citations. Let us assume that the probability that the \( n \)-th follow-up innovation cites all relevant past patents is \( \kappa^n \), where \( 0 < \kappa < 1 \). This generates a decline in the relative citation rate as a technological cluster ages. We can then derive the expected number of citations received by major as well as follow-up innovations. For major technological advances, this number is given by:

\[
\mathbb{E}[\text{cit}^M] \equiv \frac{\tau \omega}{\tau + z} \cdot 0 + \Lambda \left\{ \kappa + \Lambda \left[ \kappa^2 + \Lambda \left( \kappa^3 + \ldots \right) \right] \right\} = \sum_{j=1}^{\infty} (\kappa \Lambda)^j
\]

In the first term, \( \frac{\tau \omega}{\tau + z} \) is the probability that a successful innovation in a given product line is external and major, which creates a new cluster and does not add a citation within the existing cluster. In the second term, we have defined
\[ \Lambda \equiv \frac{\tau(1-\omega) + z}{\tau + z} \]

as the probability of a follow-up innovation, coming either from an external or an internal innovation. The probability that such an innovation yields a single citation is \( \kappa \), and therefore the probability of the \( n \)-th follow-up yielding a citation is \( (\kappa\Lambda)^n \). The expected number of citations is then the sum of all such probabilities. Since \( \kappa\Lambda < 1 \), Equation (36) can be expressed as:

\[ \mathbb{E}[\text{cit}_M] = \frac{\kappa\Lambda}{1 - \kappa\Lambda} \]

A similar derivation shows that the expected number of citations for the \( n \)-th follow-up innovation in a given technological cluster, denoted \( \mathbb{E}[\text{cit}_n^F] \), is given by:

\[ \mathbb{E}[\text{cit}_n^F] = \frac{\kappa^{n+1}\Lambda}{1 - \kappa\Lambda} \]

Therefore, \( \mathbb{E}[\text{cit}_n^F] < \mathbb{E}[\text{cit}_M] \) for any number of follow-ups, \( n \in \mathbb{N} \). Since in our baseline model with advertising smaller firms invest relatively more in external R&D, these firms also hold relatively more patents, and these patents are of higher quality on average (as measured by the number of external citations). In sum, our baseline model can generate the observed negative correlation between firm size and the fraction of top quality patents in the firm’s patent portfolio, even when there exist non-decreasing returns to scale in the R&D technology.

C.2 Alternative Advertising Functions

The main mechanisms presented in this paper rely on two important aspects of advertising. The first one is that advertising acts as a demand shifter in a way that is comparable to the effect of increased intrinsic quality from innovation. This is a necessary condition for the substitution between R&D and advertising that we obtain in our calibration exercise. Second, the existence of spillover in advertising implies that R&D intensity can be decreasing in firm size even in the absence of decreasing returns to scale in innovation. In this section, we show different alternative ways to model advertising which also lead to advertising being a demand shifter. This implies
that the results that we obtain in this paper would still qualitatively hold under the different alternative models presented below.

C.2.1 Goodwill Accumulation

In our baseline model, we assume that the advertising decision is static. Advertising expenditures affect current demand but have no long-lasting effect on consumer demand. An alternative way of modeling advertising and its effects on demand which is often used in the literature is to assume that advertising expenditures accumulate over time to increase a brand equity (so-called goodwill). Goodwill in turns acts as a demand shifter. In discrete time settings, the evolution of the stock of goodwill for good \( j \) at time \( t \) \((G_{j,t})\) is usually modeled as follows:

\[
G_{j,t} = \delta G_{j,t-1} + d_{j,t} \tag{37}
\]

where \( \delta \in [0,1] \) controls the rate at which goodwill depreciates.

Here, we show that allowing for goodwill accumulation would not change the results derived in our baseline model. We can also notice that the marketing literature shows that the depreciation rate is relatively high so that the effect of advertising on sales has almost entirely vanished after one year. In addition, the introduction of goodwill in the model only affects the decision of firms of different ages. Since the focus of our paper is on firm behavior across firm size, not modeling the evolution of goodwill over firm age is not a major concern.\(^60\)

We now present a discrete-time version of our model with goodwill and show that it does not significantly change the problem of intermediate good producers. We use Equation (37) as the evolution of goodwill over time.

The final good is now defined as:

\[
Y_t = \frac{1}{1-\beta} \int_0^1 q_{j,t}^\beta (1 + G_{j,t})^\beta y_{j,t}^{1-\beta} dj
\]

\(^60\)We can further notice that firm age does not significantly affect advertising intensity and the relative use of R&D and advertising in Tables 3 and 4.
The inverse demand function for good $j$ is given by:

$$p_{j,t} = q_{j,t}^\beta (1 + G_{j,t})^\beta y_{j,t}^{-\beta}$$

Advertising goodwill ($G_{j,t}$) is thus a demand shifter. Intermediate good producers maximize their profit subject to the inverse demand function and the dynamics of goodwill. Operational profit (profit before R&D and advertising expenditures) can be written as:

$$\pi_{j,t} = (1 - \beta)^{1-\beta} \beta \left( \frac{\bar{Q}}{w_t} \right) \frac{1-\beta}{\beta} q_j (1 + \delta G_{j,t-1} + d_{j,t})$$

which is similar to what was obtained in Section 4.5. The first order condition for advertising choice is also equivalent.

**C.2.2 Advertising in Utility**

In this section, we show that a slight modification of the baseline model which allows advertising to feature directly into the consumers’ utility function delivers a very similar allocation, and identical qualitative predictions, as the model in the main text.

There is no final good sector and the household consumes goods $j \in [0,1]$ directly. The representative household’s preferences are now represented by:

$$U = \int_0^{+\infty} e^{-\rho t} \ln (C_t) dt$$

where $C_t$ is a consumption aggregator over a mass-one continuum of quality-weighted good quantities, indexed by $j \in [0,1]$, which takes the form:

$$C_t = \frac{1}{1-\beta} \int_0^1 \bar{q}_{jt} y_{jt}^{1-\beta} dj$$

with $\beta \in (0,1)$. The flow budget constraint is, therefore:

$$\dot{A}_t = r_tA_t + w_t - \int_0^1 p_{jt} y_{jt} dj$$

where $A_0 \geq 0$ is given, and $p_{jt}$ is the price of good $j$. Each good variety $j$ is produced with
technology:

\[ y_{jt} = Q_t l_{jt} \]

where good \( j = 0 \) is the numeraire (so \( p_{0,t} = 1, \forall t \)).

Taking initial wealth \( A_0 \) as given, the representative consumer chooses a path for consumption to maximize utility subject to the flow budget constraint and the no-Ponzi condition

\[ \lim_{t \to +\infty} e^{-\int_0^t r_s ds} A_t \geq 0. \]

The optimality condition for good \( j \) yields:

\[ \omega_t p_{jt} = 1 \]

\[ C_t \left( \frac{y_{jt}}{q_{jt}} \right)^{-\beta} \]

where \( \omega_t \geq 0 \) is the Lagrange multiplier, solving

\[ \frac{\dot{\omega}_t}{\omega_t} = r_t - \rho \]

i.e., \( \omega_t = \omega_0 e^{-\int_0^t (r_s - \rho) ds} \). Recalling that \( \tilde{q}_{jt} = q_{jt} + \phi_{jt} \), Equation (38) says that the inverse demand function for goods from households firms is iso-elastic, with \( \beta \) being the price-elasticity.

Solving the incumbent firm’s problem, one can show easily that \( \omega_t = \frac{1}{Y_t} \), where \( Y_t \equiv \int_0^1 y_{jt} dj \) denotes aggregate output in the economy. Hence, the Euler equation reads \( g_t = r_t - \rho \) and the demand function becomes identical to that of the baseline model (Equation (8)). The two models are therefore qualitatively equivalent.

C.2.3 “Wasteful” Combative Advertising

In our baseline model, advertising not only shifts demand but also has an effect on consumer utility and welfare. In Section 5.3, we nevertheless show that the calibrated version of the model suggests that advertising is welfare decreasing as the level effect (the increase in how consumers value their consumption) is more than offset by the negative effect of advertising on growth through the substitution between advertising and R&D at the firm level. In this section, we show that the same results can be obtained in a model in which advertising does not increase the value of consumption in equilibrium. This can be seen as a model of combative (or predatory) advertising in which the advertising efforts of each firm (partially) cancel out in equilibrium.
Let us define the final good as:

\[ Y = \frac{1}{1 - \beta} \int_{0}^{1} q_j^\beta (1 + d_j - \iota \Phi^*)^\beta y_j^{1-\beta} \, dj \]

with \( \iota \in [0, 1] \). This implies that the effectiveness of advertising at the good level is a function of the overall level of advertising expenditures in the economy (through \( \Phi^* \), i.e. the normalized aggregate extrinsic quality in the economy). Thus, the more other firms invest in advertising, the more one firm has to invest itself in order to obtain a given return to advertising.

We obtain the following demand function:

\[ y_j = q_j (1 + d_j - \iota \Phi^*) p_j^{-\frac{1}{\beta}} \]

Intermediate good firms solve their profit maximization problem subject to this demand function and taking the overall level of advertising in the economy as given. From the firm’s perspective, advertising acts as a demand shifter in the same way as in our baseline model. Moreover, it is easy to derive \( Y_t \) in equilibrium as:

\[ Y = \left( \frac{Q}{w} \right)^{\frac{1-\beta}{\beta}} (1 - \beta)^{\frac{1-2\beta}{\beta}} [1 + (1 - \iota) \Phi^*] \bar{Q} \]

If \( \iota = 1 \), advertising has no direct effect on consumer’s utility. It can, nevertheless, have an effect on lifetime utility through its impact on the growth rate of the economy in a way that is similar to our baseline model. If \( \iota = 0 \), we return to our benchmark model.

### C.2.4 Informative Advertising

In our baseline, model advertising is purely persuasive in the sense that it shifts demand toward advertised good though increased marginal utility. Alternatively, one could consider advertising as providing relevant information about the product quality (see for instance Nelson (1974), Butters (1977), Grossman and Shapiro (1984) or Milgrom and Roberts (1986)). In this case, advertising could be socially optimal as it could reduce uncertainty or improve the quality of consumer-firm match.

In this section, we propose a simple model of informative advertising with differentiated
products. We look at a static model in which advertising is used to provide information about the quality of the goods. In particular, firms send an imperfect signal about their product quality through advertising. Consumers passively receive the information and update their prior about product quality.

Consumers maximize expected utility. We assume that the utility function is quadratic and given by:

$$U = \int_0^1 q_j y_j \, dj - a \int_0^1 q_j^2 y_j^2 \, dj$$

Before receiving signals through advertising, consumers have a prior about the quality of each good $j$. This prior is normally distributed with mean $\mu_j$ and variance $\sigma_j^2$. Through advertising, firms can send an imperfect signal ($s_j$) about their product quality, given by:

$$s_j = q_j^* + \omega_j$$

where $q_j^*$ is the actual quality of the good and $\omega_j$ is a Gaussian shock with mean 0 and variance $\sigma_\omega^2$.

We assume that higher advertising expenditures can decrease the variance of the signal. The posterior distribution of product quality (after receiving the signal) follows a normal distribution with mean and variance:

$$\mu_{post} = \frac{\mu_j / \sigma_j^2 + s_j / \sigma_\omega^2}{\sigma_j^{-2} + \sigma_\omega^{-2}}$$

$$\sigma_{post}^2 = \left[ \frac{1}{\sigma_j^2} + \frac{1}{\sigma_\omega^2} \right]^{-1}$$

The representative consumer maximizes expected utility after receiving advertising signals. The demand function for good $j$ can be written as:

$$y_j = \frac{\mu_{post} - p_j}{2a (\mu_{post}^2 + \sigma_{post}^2)}$$

For simplicity, let us assume the limiting case as $\sigma_j$ goes to infinity (a diffuse prior). In this
case, the posterior distribution has mean $\mu_{post} = s_j$ and variance $\sigma^2_{post} = \sigma^2_\omega$. Therefore, the demand for any good $j$ is a decreasing function of the variance of the signal. Since the precision of the signal is increasing in advertising expenditures, advertising acts as a demand shifter as in our baseline model.

C.2.5 Advertising and the Price-Elasticity of Demand

A strand of the advertising literature has focused on the effect of advertising on the elasticity of demand (see, for instance, Molinari and Turino (2009) and Benhabib and Bisin (2002)). In this section, we present a model in which advertising can change the price-elasticity of demand. We further show that the demand shifting property of advertising is maintained so that the results from our baseline model could be obtained in such a framework as well.

Following Molinari and Turino (2009), we write:

$$Y = \left[ \int_0^1 q_j^\frac{1}{\epsilon} (y_j + D(d_j))^{-\frac{1}{\epsilon}} \, dj \right]^{\epsilon^{-1}}$$

where $D$ is a decreasing function, with $D(0) \geq 0$. The inverse demand function can be written as:

$$p_j = (q_j Y)^{1/\epsilon} [y_j + D(d_j)]^{-1/\epsilon}$$

Setting $D'(d_j) < 0$ as in Molinari and Turino (2009), we obtain that advertising acts as a positive demand shifter. Furthermore, the price-elasticity of demand is equal to:

$$\left| \frac{\partial y_j / y_j}{\partial p_j / p_j} \right| = \epsilon \left( 1 + \frac{D(d_j)}{y_j} \right)$$

Thus, advertising decreases the price-elasticity of demand, ceteris paribus. Intuitively, by conducting advertising, firms alter the substitutability between goods, and make them more price-inelastic.
C.2.6 Using a CES Production Function

We assume that the production function for the final good is the following CES function:

\[
Y = \left[ \int_0^1 q_j^{\frac{1}{\epsilon}} (1 + d_j)^{\frac{1}{\epsilon} y_j^{\epsilon - 1}} \right]^{\frac{1}{1 + \epsilon}} d_j
\]

The demand function for intermediate goods can be obtained from the first order condition of the final good producer and be written as:

\[
y_j = \frac{Y q_j (1 + d_j)}{p_j^{\epsilon}} \quad (39)
\]

Intermediate good producers then maximize their objective function subject to Equation (39) and taking \( Y \) as given. Advertising acts as a demand shifter and the problem of intermediate good producers is similar to the one in Section 4.5.