Advertising, Innovation and Economic Growth

Online Appendix: Not For Publication

Laurent Cavenaile* Pau Roldan-Blanco†

April 2020

Contents

*UTSC and Rotman School of Management, University of Toronto, 105 Saint George Street, Toronto, Ontario, M5S 3E6, Canada. Email: `laurent.cavenaile@utoronto.ca`

†Bank of Spain, Alcalá 48, Madrid, 28014, Spain. Email: `pau.roldan@bde.es`

A Comparative Statics

In this Appendix:

1. We show that our model can be solved with increasing returns to scale in R&D technology.
2. We show that the model can deliver both substitutability and complementarity between R&D and advertising.
3. We show which parameters drive the substitution/complementarity margin in the calibrated model.

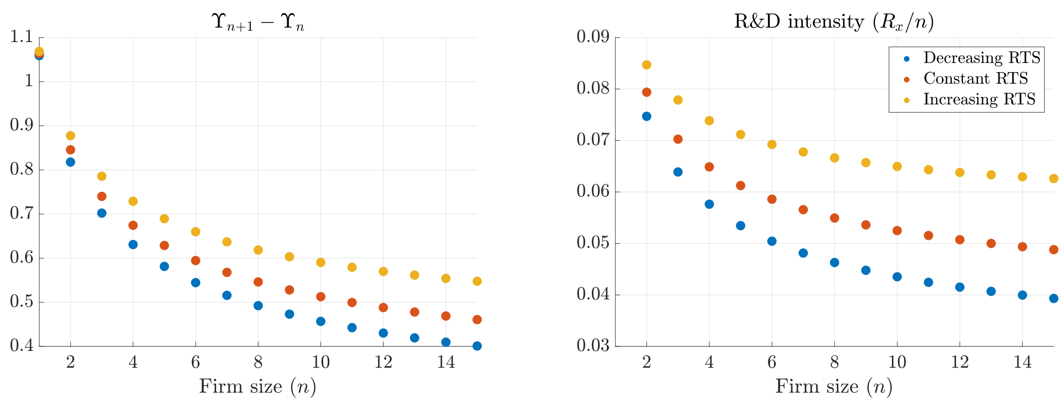

Figure A.1: Comparative statics for R&D intensity across n, for $\sigma + \widetilde{\psi} > 1$ (decreasing returns), $\sigma + \widetilde{\psi} = 1$ (constant returns), and $\sigma + \widetilde{\psi} < 1$ (increasing returns).

First, 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 A.1 we show three solutions of the model, using decreasing, constant, and increasing returns to scale in external R&D, respectively. The left panel shows the marginal value from acquiring an additional product ($\Upsilon_{n+1} - \Upsilon_n$), and the right panel shows the optimal R&D intensity, as functions of firm size. 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 is steeper). Symmetrically, having increasing returns in innovation increases optimal intensity for all n. Interestingly, so long as the degree of increasing returns is not too strong, smaller firms may still optimally choose to invest relatively more in R&D in order to reap the benefits of advertising spillovers, as is the case in the figure. This occurs even though the marginal benefit from innovation is lower than in the other two cases (so that the R_x/n line flattens but it is still decreasing).

Next, we show that the model can deliver both substitutability and complementarity between R&D and advertising. Figure A.2 depicts, for a numerical example, the change in external R&D across firm size for different levels of advertising efficiency as well as total incumbent R&D and

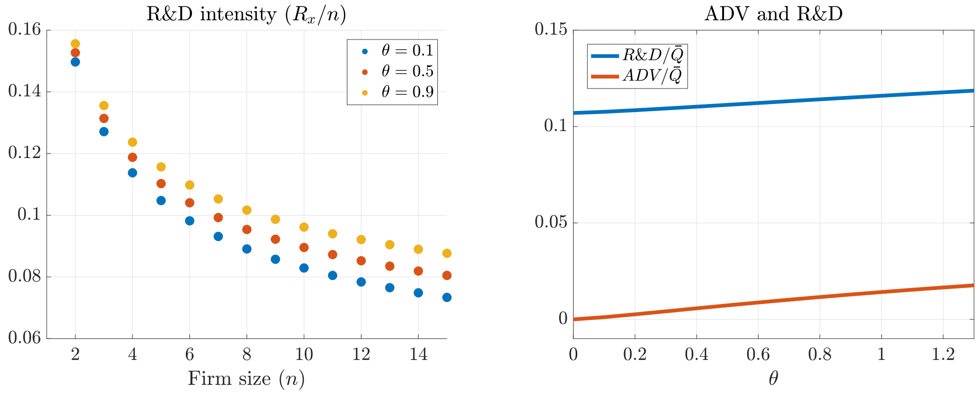

Figure A.2: External R&D intensity in BGP (left panel), and R&D and advertising expenditures in BGP, for different values of advertising efficiency, θ (right panel). All variables are normalized by $\bar{Q}$.

advertising.[1] In this example, and contrary to the calibrated model, R&D and advertising are complements: an increase in advertising efficiency is associated with larger investment in both advertising and R&D.

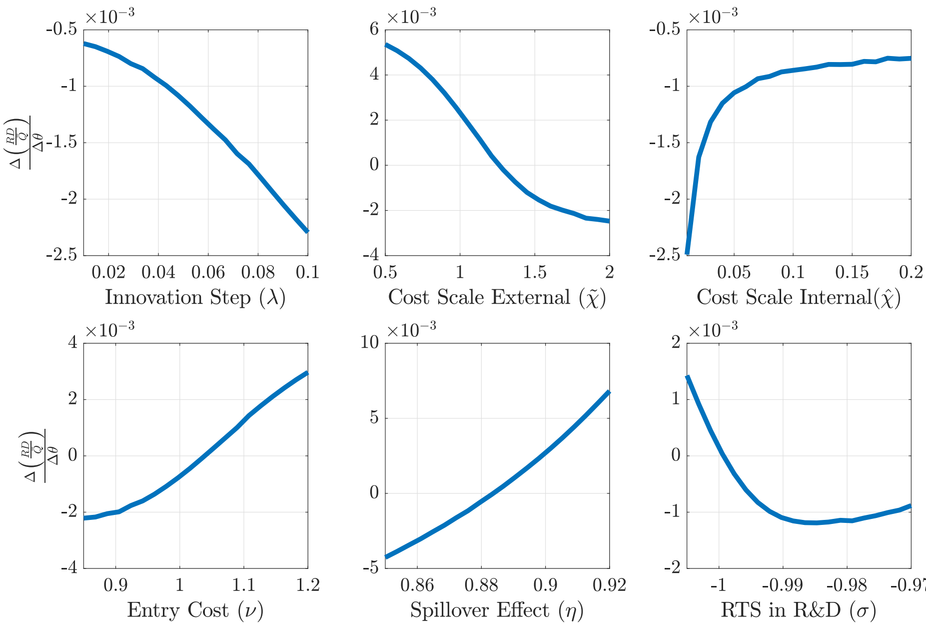

Figure A.3: Change in incumbent R&D, for a small change in θ, evaluated at the calibrated θ. The figure shows this across different parameters, with all the remaining parameters being fixed at their calibrated values from Table 3.

To understand which parameters drive this margin in the calibrated version of the model, in Figure A.3 we plot the change in incumbent R&D (normalized by $\bar{Q}$) for a small change in θ away from its calibrated value (a proxy for the derivative of R&D with respect to θ). We

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

plot this change for different values of each one of the internally calibrated parameters in some neighborhood around their calibrated value, where we fix all the remaining parameters to their calibrated values (from Table 3). When this derivative is negative, an increase in θ (i.e. a decrease in the cost of advertising) leads to a decrease in R&D, so R&D and advertising are substitutes. When the derivative is positive, R&D and advertising are complements.

The figure shows that the key parameters that control for this margin are the cost scale parameter in the external R&D technology ($\widetilde{\chi}$), the entry cost (ν), the advertising spillover parameter (η), and the parameter controlling the degree of returns to scale in R&D (σ). Note, however, that we obtain substitution for all values of the innovation step (λ), as well as the cost scale parameter in the internal R&D technology ($\widehat{\chi}$). Intuitively, the parameters related to the external R&D technology are drivers of the substitution/complementarity margin because the R&D-advertising interaction in the model operates through external innovations. In particular, we obtain that the model delivers complementarity when external R&D is relatively cheaper (low enough values of $\widetilde{\chi}$), and when it scales weakly with firm size (low enough values of σ). Additionally, as firm entry in the model occurs through external R&D, we also obtain complementarity when entry is relatively cheaper (low enough values of ν). Finally, a high enough spillover in advertising (η) also generates complementarity, everything else equal.

B The BGP Algorithm

The following describes the steps of the BGP algorithm:

1. Guess a number $\Phi^* > 0$.
2. Compute Υ_1 by imposing $\Upsilon_2 = 2\Upsilon_1$ in Equation (12) at $n = 1$.
   (a) Compute g from (8), τ implied by (9), Γ from (11), and z from (13a).
   (b) Compute $\{\Upsilon_n\}_{n=2}^{+\infty}$ iterating (12) forward from Υ_1 and $\Upsilon_0 = 0$, $\{x_n\}_{n=1}^{+\infty}$ from (13b), and F/x_e using that $\sum_{n=1}^{N} \mu_n = 1$, where μ_n comes from (10).
3. Verify convergence of the firm size distribution. If there is no convergence by iteration $k \in \mathbb{N}$, go back to step 2 with the new guess for Υ_1 equal to the solution of (12) at $n = 1$ when $\Upsilon_2 = 2\Upsilon_1 - \varepsilon_k$, for a small $\varepsilon_k > 0$.
4. Compute Φ^* as the solution to (5), 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 Υ_1 such that Υ_n can be weakly concave (i.e, $(\Upsilon_{n+1} - \Upsilon_n)$ decreasing). In particular, we force Υ_n to be straight line from $n = 1$ to $n = 2$ (note

$\Upsilon_2 - \Upsilon_1 = \Upsilon_1 - \Upsilon_0$). If μ_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 Υ_1 was incorrect. Then, we iterate on new guesses for Υ_1 to allow for more concavity on the Υ_n sequence (indeed, note that in any iteration $k \geq 1$ we always start the Υ_n sequence at a Υ_1 such that $\Upsilon_2 - \Upsilon_1 < \Upsilon_1 - \Upsilon_0$). In step 3, we 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 Φ^* guess in case of no convergence, and use the resulting fixed-point as the new guess.

C Model Extensions

C.1 Patents and Major Innovations

Besides differences in R&D intensity across size, 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. Two key findings are noteworthy:

Fact 1. 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).

Fact 2. These firms also tend to patent relatively more on average.[2]

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.

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 (2018), 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.

[2] Both facts have been documented by Akcigit and Kerr (2018). On Fact 1, they show that the fraction of a firm's patent in the top patent quality decile is decreasing with firm size. On fact 2, they show that there is a negative correlation between firm size (measured by employment) and patent intensity (measured as the ratio of patents to employment).

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 ω 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:

$$q_{j,t+\Delta t} = q_{jt} + \begin{cases} \lambda^H q_{jt} & \text{w.prob. } \omega\tau_t\Delta t + o(\Delta t) & \text{[External, major advance]} \\ \lambda^L q_{jt} & \text{w.prob. } (1-\omega)\tau_t\Delta t + o(\Delta t) & \text{[External, follow-up]} \\ \lambda^I q_{jt} & \text{w.prob. } z_{jt}\Delta t + o(\Delta t) & \text{[Internal]} \\ 0 & \text{w.prob. } 1 - \tau_t\Delta t - z_{jt}\Delta t - o(\Delta t) & \text{[No innovations]} \end{cases}$$

Therefore, the average quality improvement from a successful external innovation is equal to $\lambda^E \equiv \omega\lambda^H + (1-\omega)\lambda^L$. The growth and creative destruction rates of the economy are still given by (8) and (9), respectively. Additionally, the model now delivers Fact 1: smaller firms, while exhibiting a higher R&D intensity because of the advertising spillover effect, also produce relatively more major innovations.

Next, we show that the extended 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 κ^n, where $0 < \kappa < 1$. This generates a decline in the relative citation rate as a technological cluster ages. The expected number of citations received by major as well as follow-up innovations for major technological advances is:

$$\mathbb{E}[cit^M] \equiv \underbrace{\frac{\tau\omega}{\tau+z}}_{\substack{\text{External and major}\\ \text{innovation}}} \cdot 0 \quad + \underbrace{\Lambda\left\{\kappa + \Lambda\left[\kappa^2 + \Lambda\left(\kappa^3 + \dots\right)\right]\right\}}_{\text{Follow-up innovations}} = \sum_{j=1}^{\infty} (\kappa\Lambda)^j \qquad \text{(C.1)}$$

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 κ, 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 (C.1) can be expressed as:

$$\mathbb{E}[cit^M] = \frac{\kappa\Lambda}{1-\kappa\Lambda}$$

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

$$\mathbb{E}[cit_n^F] = \frac{\kappa^{n+1}\Lambda}{1-\kappa\Lambda}$$

Therefore, $\mathbb{E}[cit_n^F] < \mathbb{E}[cit^M]$ for any number of follow-ups, $n \in \mathbb{N}$. Since in the 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 model with advertising 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 mechanism presented in the paper relies on one key assumption: 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 (though not sufficient) condition for the substitution between R&D and advertising that we obtain in our calibration.

The purpose of this section is to propose different alternative ways to model advertising, and show that they also lead to advertising being a demand shifter. Thus, the results obtained in this paper would 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. The evolution of goodwill is:

$$\dot{G}_j = d_j - \delta G_j \tag{C.2}$$

where $\delta \in [0,1]$ controls the rate at which goodwill depreciates. Allowing for goodwill accumulation would not qualitatively change the results derived in our baseline model. This said, the marketing literature shows that the depreciation rate is relatively high, with the effect of goodwill on sales almost entirely vanishing after one year (see for instance Assmus *et al.* (1984) and Dubé *et al.* (2005)). With our calibration at the yearly frequency, modeling advertising as a static decision is not likely to be a major concern. 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 for our purposes.[3]

In the goodwill model, the final good is now defined as:

$$Y_t = \frac{1}{1-\beta}\int_0^1 q_{j,t}^{\beta}(1+G_{j,t})^{\beta} y_{j,t}^{1-\beta}\ \mathrm{d}j$$

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.

C.2.2 Advertising in Utility

Next, 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 baseline model. In this extension, 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:

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

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:

$$C_t = \frac{1}{1-\beta}\int_0^1 \widetilde{q}_{jt}^{\beta} y_{jt}^{1-\beta}\mathrm{d}j$$

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

[3] Moreover, firm age does not significantly affect advertising intensity and the relative use of R&D and advertising, as seen in Table 1.

$$\dot{A}_t = r_t A_t + w_t - \int_0^1 p_{jt} y_{jt} \mathrm{d}j$$

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

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

where good $j = 0$ is the numeraire (so $p_{0,t} = 1$, $\forall t$). The optimality condition for good j yields:

$$\omega_t p_{jt} = \frac{1}{C_t} \left(\frac{y_{jt}}{\widetilde{q}_{jt}} \right)^{-\beta}$$

where $\omega_t \geq 0$ is the Lagrange multiplier, solving $-\frac{\dot{\omega}_t}{\omega_t} = r_t - \rho$. Recalling that $\widetilde{q}_{jt} = q_{jt} + \phi_{jt}$, we have that the inverse demand function for goods from households firms is iso-elastic, with β being the price-elasticity. Solving the incumbent firm's problem, one can show easily that $\omega_t = \frac{1}{Y_t}$, where $Y_t \equiv \int_0^1 y_{jt} \mathrm{d}j$ 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. 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. Nevertheless, in Appendix C we showed that in the calibrated version of the model 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.

Next, 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 technology as:

$$Y = \frac{1}{1-\beta} \int_0^1 q_j^\beta (1 + d_j - \iota \Phi^*)^\beta y_j^{1-\beta} \ \mathrm{d}j$$

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 Φ^*, 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:

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

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 goods through 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 to illustrate how advertising can act as a demand shifter. 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. The utility function is quadratic and given by:

$$U = \int_0^1 q_j y_j \, \mathrm{d}j - \frac{\alpha}{2}\int_0^1 q_j^2 y_j^2 \, \mathrm{d}j$$

Before receiving signals through advertising, consumers have a prior about the quality of each good j. This prior is normally distributed with mean μ_j and variance σ_j^2. Through advertising, firms can send an imperfect public 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 ω_j is a Gaussian shock with mean zero and variance σ_ω^2. As in other models of informative advertising (e.g. Erdem *et al.* (2008)), we assume that higher advertising expenditures can decrease the variance of the signal.

It is not difficult to show that the posterior distribution of product quality (after receiving the signal) follows a normal distribution with mean and variance:

$$\begin{aligned}\mu_{post} &= \frac{\mu_j/\sigma_j^2 + s_j/\sigma_\omega^2}{\sigma_j^{-2}+\sigma_\omega^{-2}} \\ \sigma_{post}^2 &= \left(\frac{1}{\sigma_j^2}+\frac{1}{\sigma_\omega^2}\right)^{-1}\end{aligned}$$

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

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

Note that:

$$\frac{\partial \mu_{post}}{\partial \sigma_\omega^2} = -\left(\frac{\sigma_j^2}{\sigma_j^2+\sigma_\omega^2}\right)^2 \frac{s_j-\mu_j}{\sigma_j^2} \quad \text{and} \quad \frac{\partial \sigma_{post}^2}{\partial \sigma_\omega^2} = \left(\frac{\sigma_j^2}{\sigma_j^2+\sigma_\omega^2}\right)^2$$

This means that, by doing advertising (which lowers σ_ω^2), firms can effectively lower the posterior variance of households. Furthermore, if the signal is sufficiently optimistic relative to the household's prior (i.e. $s_j > \mu_j$), increasing advertising expenditure also increases the posterior mean. To see how this affects the final demand for the good, note:

$$\frac{\partial y_j}{\partial \sigma_\omega^2} = -\frac{\partial \sigma_{post}^2}{\partial \sigma_\omega^2}\left(\frac{\frac{s_j-\mu_j}{\sigma_j^2}(p_j+\sigma_{post}^2)+(\mu_{post}-p_j)}{\alpha(\mu_{post}^2+\sigma_{post}^2)^2}\right)$$

When the signal is optimistic ($s_j > \mu_j$), advertising will unambiguously shift demand, as it helps household reduce the posterior variance and upgrade the posterior mean, both of which boost demand. When the signal is pessimistic ($s_j < \mu_j$), there is a trade-off: increased advertising helps reduce the posterior variance, but it also lowers the posterior mean, so the net effect depends on fundamentals. For example, when the prior is diffuse ($\sigma_j \to +\infty$), the posterior distribution mean and variance satisfy $\mu_{post} \to s_j$ and $\sigma_{post}^2 \to \sigma_\omega^2$, respectively, and advertising expenditures unambiguously shift demand. More generally, for the case of pessimistic signals, there always exists a prior variance that is diffused enough for advertising to shift demand. In these cases, 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 (2015) 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 (2015), we write:

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

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

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

Setting $D'(d_j) < 0$ as in Molinari and Turino (2015), 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.

D Description of R&D Tax Credit Measures

This section describes the measures of R&D tax credits which are used in our empirical investigation of the substitution between R&D and advertising in Section III.C. Besides the statutory credit rate, we use the following measures:

- First, the tax-adjusted credit rate 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 s at time t as:

$$\text{Tax-adjusted Rate}_{st} = \text{Statutory Credit Rate}_{st} \times (1 - \sigma_{st} \times \text{Tax Rate}_{st})$$

  where σ_{st} 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.

- Second, we use an alternative measure of the marginal effective R&D tax credit, proposed by Wilson (2009). This measure is available only until 2006. It acknowledges the different definitions of the R&D expenditures which are eligible for 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, in turn, may be a moving average of past R&D expenditures. For such states, the moving-average base is usually computed as the product of firm sales and the R&D-to-sales ratio over the n previous periods. For a firm with R&D expenditures above the base level, the marginal effective tax credit rate (m_{st}) is computed as:

$$m_{st} = \text{Statutory Credit Rate}_{st} \times (1 - \sigma_{st} \times \tau^e_{st}) \times \left(1 - \frac{1}{n}\sum_{k=1}^{n}(1 + r_{t+k})^{-k}\right)$$

  where r is the real interest rate, n is the number of periods over which the moving-average base is calculated, and τ^e_{st} 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. The rate also takes into account the fact that R&D tax credit are subject to corporate taxation in some states.

  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 these cases, the marginal effective credit rate is equal to the (tax-adjusted) statutory credit rate.

- 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 *et al.* (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:

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

  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 δ is the depreciation rate of R&D capital. Following Wilson (2009), we set $v = 0.5$, $z = 0.525$ and $\delta = 0.15$.

References

Akcigit, U. and Kerr, W. R. (2018). Growth through heterogeneous innovations. *Journal of Political Economy*, **126** (4), 1374–1443.

Assmus, G., Farley, J. U. and Lehmann, D. R. (1984). How advertising affects sales: Meta-analysis of econometric results. *Journal of Marketing Research*, **21** (1), 65 – 74.

Benhabib, J. and Bisin, A. (2002). Advertising, mass consumption and capitalism. *Working Paper, New York University*.

Bloom, N., Griffith, R. and Van Reenen, J. (2002). Do R&D tax credits work? Evidence from a panel of countries 1979 - 1997. *Journal of Public Economics*, **85** (1), 1 – 31.

Butters, G. R. (1977). Equilibrium distributions of sales and advertising prices. *The Review of Economic Studies*, **44** (3), 465 – 491.

Dubé, J.-P., Hitsch, G. J. and Manchanda, P. (2005). An empirical model of advertising dynamics. *Quantitative Marketing and Economics*, **3** (2), 107 – 144.

Erdem, T., Keane, M. and Sun, B. (2008). A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science*, **27** (6), 949 – 1136.

Grossman, G. M. and Shapiro, C. (1984). Informative advertising with differentiated products. *The Review of Economic Studies*, **51** (1), 63 – 81.

Hall, R. E. and Jorgenson, D. W. (1967). Tax policy and investment behavior. *The American Economic Review*, **57** (3), 391 – 414.

Milgrom, P. and Roberts, J. (1986). Price and advertising signals of product quality. *Journal of Political Economy*, **94** (4), 796 – 821.

Molinari, B. and Turino, F. (2015). Advertising and business cycle fluctuations. *Working Paper, Universitat Pompeu Fabra*.

Nelson, P. (1974). Advertising as information. *Journal of Political Economy*, **82** (4), 729 – 754.

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