ARTICLE IN PRESS

Journal of Monetary Economics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jme

Style over substance? Advertising, innovation, and endogenous market structure☆

Laurent Cavenaile[a], Murat Alp Celik[b,*], Pau Roldan-Blanco[c,d], Xu Tian[e]

[a] *University of Toronto Scarborough and Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON, M5S 3E6, Canada*
[b] *University of Toronto, 150 St. George Street, Toronto, ON, M5S 3G7, Canada*
[c] *Universitat Autònoma de Barcelona, UAB Campus, Edifici B, 08193 Bellaterra, Barcelona, Spain*
[d] *Barcelona School of Economics, Spain*
[e] *Terry College of Business, University of Georgia, 620 South Lumpkin Street, Athens, GA, 30602, USA*

ARTICLE INFO

JEL classification:
E20
M30
O30
O40

Keywords:
Innovation
Advertising
Markups
Growth
Industry dynamics
Misallocation
Business dynamism

ABSTRACT

While firms use both innovation and advertising to boost profits, markups, and market shares, their broader social implications vary substantially. We study their interaction and analyze their implications for competition, industry dynamics, growth, and welfare. We develop an oligopolistic general-equilibrium growth model with firm heterogeneity. Market structure is endogenous, and firms' production, innovation, and advertising decisions interact strategically. We find advertising reduces static misallocation, but also depresses growth through a substitution effect with R&D. Although advertising is found to be socially useful, taxing it could simultaneously increase dynamic efficiency, contain excessive advertising spending, and raise revenue, while still reducing misallocation.

1. Introduction

Firms compete against each other in a variety of ways to increase profits and in pursuit of a dominant position in their industry. A well-known tool at the firms' disposal is innovation. Firms spend resources on research and development (R&D) to develop new products or more efficient technologies, thereby expanding into new product markets or raising productivity to cut costs and increase sales relative to their competitors. Advertising is the other most prominent intangible investment through which firms can achieve

☆ We would like to thank our editors Boragan Aruoba and Felipe Saffie, our anonymous referees, as well as David Argente, Salome Baslandze, Maarten De Ridder, James Fenske, Javier Fernandez, Jeremy Greenwood, Douglas Hanley, Giammario Impullitti, Charles Jones, Michael Klein, Omar Licandro, Yueyuan Ma, Atif Mian, Sara Moreira, Pedro Monteiro, Sergio Ocampo, Miquel Oliver i Vert, Guillermo Ordoñez, Jesse Perla, Thomas Philippon, Michael Peters, Nicolas Sahuguet, Kenneth Simons, Christopher Tonetti, Gustavo Ventura, Nicolas Vincent, Toni Whited, Peifan Wu, Daniel Yi Xu, Mehmet Yorukoglu, and conference and seminar participants at Banco de España, Nottingham, FMA, NFA, AFA, CMSG, UAB, Rensselaer, EWMES, HEC Montreal, BSE Summer Forum, SED, CEA, EARIE, SEA, UQAM, Bristol, Laval, Zurich, Bank of Canada, ASU, Pittsburgh/CMU, Richmond Fed, and UPenn for their helpful comments. Cavenaile and Celik gratefully acknowledge financial support from SSHRC Insight Development Grants (517148 and 503178). Tian gratefully acknowledges financial support from the Business, Systems, and Technology Innovation Research Grant from the Terry College of Business at the University of Georgia.

* Corresponding author.
E-mail addresses: laurent.cavenaile@utoronto.ca (L. Cavenaile), murat.celik@utoronto.ca (M.A. Celik), pau.roldan@barcelonagse.eu (P. Roldan-Blanco), xu.tian@uga.edu (X. Tian).

https://doi.org/10.1016/j.jmoneco.2024.103683
Received 29 November 2023; Received in revised form 2 September 2024; Accepted 4 September 2024
Available online 10 September 2024

Please cite this article as: Laurent Cavenaile et al., *Journal of Monetary Economics* , https://doi.org/10.1016/j.jmoneco.2024.103683

the same desired outcomes, by shifting demand toward their products and away from those of their rivals. In fact, as a society, we spend as much on developing new products and technologies ("substance") as on marketing them ("style").[1]

Since both activities pursue similar goals, firms' decisions to innovate and to advertise inexorably interact — within the firm itself, as well as across all the firms in the same industry. Yet, the aggregate consequences of these two investments are dramatically different: while innovation is recognized as an essential engine of economic growth, advertising does not directly contribute to long-run productivity. Moreover, under intense competition, a "rat race" in advertising can lead to socially excessive spending.

In this paper, we explore and quantify the aggregate welfare consequences of advertising when both innovation and advertising are used strategically by firms to compete with one another. To do so, we build a model of firm and industry dynamics which can elucidate the role of innovation and advertising for market concentration, markups, and productivity growth, offering a realistic representation of how these two intangible investments interact and relate to competition at both the micro and the aggregate levels. Our key finding is that advertising plays an allocative role which is critical for welfare, both statically and dynamically. Dynamically, we find it to be a substitute to innovation, reducing productivity growth. Statically, advertising helps reallocate resources toward the most productive firms, which generates welfare gains from reduced misallocation of physical inputs. In net, these static gains are sufficiently large to quantitatively outweigh the dynamic welfare losses. There is room for policy intervention in the form of advertising taxes, which we investigate.

We start our analysis by exploring the micro-data. Using balance-sheet information for publicly listed companies in the United States, we uncover two robust empirical regularities. First, there exists a non-monotonic relationship between innovation, advertising, and market share at the firm level. In particular, and new to the literature, we document a hump-shaped relationship between the level of advertising and a firm's relative sales within its industry: advertising spending is on average highest for firms that are neither too large nor too small within their industry. Second, a firm's market share is positively associated with its own advertising, but negatively associated with the advertising efforts of its competitors, hinting at a strategic component in advertising.

Motivated by these facts, we build a new endogenous growth model, which, unlike previous models in the literature, can replicate the aforementioned empirical regularities. In this model, firms oligopolistically compete by making production, innovation, and advertising decisions. Advertising can be used by firms to shift demand towards their products, and thereby boost their market shares and markups, which in turn indirectly affects incentives to innovate. We model the economy as a continuum of industries, each populated by endogenously determined numbers of small and large firms. Large (so-called "superstar") firms behave strategically, aware of the effects of their production, advertising, and innovation decisions on other firms in their industry. Small (so-called "fringe") firms, by contrast, are infinitesimally small, charge zero markups, and make no advertising decisions, but can innovate to come up with a breakthrough innovation and join the group of large firms.

R&D and advertising are modeled as intangible expenditures which can improve a firm's market share through different channels. R&D is modeled following the tradition of the step-by-step innovation literature, in which successful innovation improves the firm's productivity. We model advertising as a demand shifter, and akin to a zero-sum game: advertising expenditures increase the perceived quality of the firm's product, making it more appealing to consumers, but also lower the perceived quality of all the competitors' products.[2] Large firms, which are heterogeneous in productivity, choose advertising optimally to maximize static profits, taking into account the effects of their choices on their own market share. In equilibrium, the differential use of advertising across firms can magnify productivity differences and have quantitatively significant implications for within-industry markup dispersion and, as a consequence, allocative efficiency. Moreover, because in equilibrium firms are heterogeneous in their use of advertising, there is a dynamic interplay between advertising and R&D decisions, which, at the aggregate level, has an impact on the rate of economic growth, business dynamism, and social welfare.

To quantify these various channels, we calibrate the parameters of the model to fit key macroeconomic aggregates from the United States, as well as empirical patterns relating to advertising and innovation at the micro level, specifically the non-monotonic relationships between innovation, advertising, and market share that we document in our empirical analysis. Our calibrated model matches these targeted moments well. It also delivers predictions that are in line with some other relevant (and untargeted) features of the data, such as industry concentration ratios, firm-level correlations between advertising, R&D, and markups, and the relationship between a firm's market share, its own advertising, and the advertising efforts of its competitors.

Using the calibrated model, we run a series of counterfactual exercises to disentangle the growth and welfare effects of advertising through its interplay with R&D, markups, and market structure. We start by comparing the calibrated baseline economy with a counterfactual economy where advertising is shut down completely (e.g., it is prohibitively expensive for firms). This simple experiment reveals the main effects of advertising in our economy. First, shutting down advertising increases firm-level investment in R&D, by both large and small firms, which increases aggregate innovation as well as the rate of economic growth.[3] Second, shutting down advertising affects markups and allocative efficiency through changes in the competitive structure of industries. The effects along this margin are theoretically ambiguous. On the one hand, markups decrease on average as large firms cease to use advertising as a tool to shift demand and profits toward their products. This spurs firm entry and boosts business dynamism. On the

[1] Since 1980, R&D accounts for 2.44% of GDP in the US, whereas advertising alone represents 2.20%. The figures for advertising do not include in-house firm expenses related to sales, which would increase the fraction of resources devoted to marketing further (found to be around 7–8% by Arkolakis (2010)).

[2] In a model extension in Section 6, we relax this assumption by varying the degree of combativeness in advertising. A re-calibration of this more general model suggests that the degree of combativeness is very close to what we assume in the baseline model.

[3] Thus, advertising and R&D are substitutes in our calibrated economy, consistent with the empirical findings in Cavenaile and Roldan-Blanco (2021). Under different parameter values, however, advertising and R&D could be complements rather than substitutes. Therefore, our substitution result is a quantitative finding rather than a theoretical implication of the framework.

other hand, advertising improves allocative efficiency. While increasing markups, advertising reallocates physical inputs away from the less efficient firms and towards the more efficient industry leaders, simultaneously amplifying the relative perceived quality of the more abundant and cheaper-to-produce varieties. While markups themselves lower efficiency, the latter two effects quantitatively dominate. In net, advertising substantially improves static allocative efficiency.[4]

Though desirable statically through its reallocation properties, advertising is also diverting resources away from innovation, leading to dynamic welfare losses through depressed economic growth. What is the net welfare effect? To answer this question, we analytically decompose the change in consumption-equivalent welfare between balanced growth path equilibria.[5] We find substantial quantitative differences between static and dynamic welfare changes. Statically (i.e., without adjustments in the industry state distribution), shutting down advertising results in a welfare loss of 3.64% in consumption-equivalent terms, mostly coming from the aforementioned losses in allocative efficiency. Taking dynamic aspects into consideration undoes some of these losses in the long run, as shutting down advertising also changes the distribution over industry states and increases the rate of economic growth through the substitution effect between R&D and advertising. Yet, on the net, the static reallocation forces prevail: shutting down advertising results in a welfare loss, albeit a small one (0.86%), despite the potentially wasteful (zero-sum game) nature of advertising in our baseline model. In a series of extensions, we find that the net positive role of advertising for welfare is in fact a robust feature of our quantitative model.[6]

In light of these results, we consider the implications for policy intervention. Although we conclude that shutting down advertising would reduce welfare, we find that advertising should be taxed, rather than subsidized. How does one reconcile the two findings? Higher taxes on advertising expenses discourage firms from investing resources in advertising, resulting in both direct gains in the consumption-to-output ratio, and indirect gains from improved incentives for innovation and growth. However, the taxes do not cause as large a drop in static allocative efficiency as a complete shutdown would: while the overall spending on advertising declines, more productive superstars still continue to spend more on advertising than less productive ones. Therefore, the positive effects of advertising in reducing static misallocation are still present, even under high tax rates. In other words, an advertising tax reduces the excessive spending on advertising due to the "rat race" between the superstars, while still largely preserving the distribution of market shares in equilibrium. This makes advertising an ideal candidate for taxation to raise revenues while simultaneously increasing dynamic efficiency. Considering (i) advertising goes untaxed in most countries with few exceptions, (ii) the disproportionate market share of tech giants such as Google and Meta in digital advertising, and (iii) the recent proposals by policymakers to change this status quo, by how much we should tax advertising is a timely and significant question, which we try to address.[7]

Literature review. Our paper is primarily related to the literature that studies the implications of intangible investments, in the form of advertising and customer capital, for firm, industry, and macroeconomic dynamics (e.g., Dinlersoz and Yorukoglu (2012), Gourio and Rudanko (2014), Molinari and Turino (2017), Crouzet and Eberly (2019), Argente et al. (2023), Einav et al. (2022), Ignaszak and Sedláček (2022), Pearce and Wu (2022), Dinlersoz et al. (2023), Cavenaile et al. (2023), and Greenwood et al. (2024)). The literature has investigated, for instance, how intangibles may be behind several trends related to business dynamism, market concentration, and markups (e.g., Cavenaile et al. (2019), Weiss (2020), Feijoo Moreira (2021), Aghion et al. (2023), and De Ridder (2024), or how they may affect markup cyclicality (e.g., Roldan-Blanco and Gilbukh (2021)), the transmission channels of monetary policy (e.g., Morlacco and Zeke (2021)), and the behavior of exporters and international prices (e.g., Drozd and Nosal (2012) and Fitzgerald et al. (2023)).

In this literature, our model is most closely related to recent macroeconomic models with advertising such as Afrouzi et al. (2023), Cavenaile and Roldan-Blanco (2021), Rachel (2022), Baslandze et al. (2023), Cavenaile et al. (2023), Klein and Şener (2023), and Greenwood et al. (2024). Our paper contributes to this body of work by modeling advertising decisions at the firm level as directly affecting market shares and markups, as suggested by our empirical findings. Unlike previous papers in this literature, our framework, which builds on Cavenaile et al. (2019), allows us to study how the interaction between R&D and advertising affects market concentration and markups, and through this channel it highlights a new reallocative role for advertising that turns out to be quantitatively essential for the welfare effects of advertising.

Our paper also contributes to a long tradition of modeling advertising in economics and finance (e.g., Dorfman and Steiner (1954), Butters (1977), Becker and Murphy (1993), Benhabib and Bisin (2002)).[8] In the literature, advertising is commonly modeled as a demand shifter. Following this tradition, we model advertising as a technology that shifts consumer preferences for certain goods

[4] We should highlight that, although advertising improves static allocative efficiency on the net, this does not mean that individual firms do not over- or under-advertise. The solution to the static social planner's problem in Section E.2 shows that, generically, firms pick socially sub-optimal advertising rates.

[5] While our welfare change results are necessarily normative, all our findings regarding innovation, growth, business dynamism, and input misallocation, among other things, are positive results that do not hinge on how advertising is treated in welfare calculations. In Section 6, we consider two alternatives for normative implications.

[6] Our two main model extensions change the effect of advertising on consumer preferences. In one extension, we assume that advertising is partly deceptive, in that it does not increase consumer utility ex-post in a way that is consistent with consumers' ex-ante demands. In the other, we vary the degree of advertising combativeness, i.e., of the negative impact of a firm's advertising efforts on the perceived quality of other firms' products. In all of our quantitative experiments, advertising is still found to be socially desirable.

[7] Recently, the state of Maryland has sought to tax digital advertising revenues, but it was struck down by a Circuit Court. Several US states and Canada are also considering imposing similar taxes, but only on digital advertising. Nevertheless, digital advertising constitutes more than 64% of all advertising in North America as of 2021, which is projected to increase further.

[8] Bagwell (2007) provides a comprehensive survey of this literature.

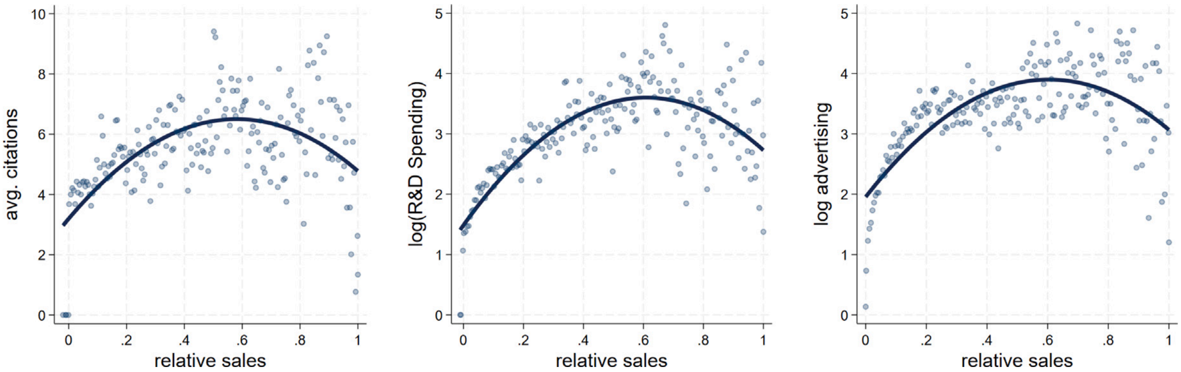


Fig. 1. Innovation, R&D spending, advertising, and firm relative sales.
Notes: From left to right, this figure displays the relationship between the raw values of average patent citations, log R&D expenses, and log advertising expenses, and firm relative sales. We divide the relative sales into 200 quantiles and calculate the average values for each quantile. The blue curve is a quadratic fit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to the detriment of competitors' products, which in equilibrium means that firms can use advertising to shift demand toward their own goods. To draw this connection between firm-level advertising and market structure, we rely on observations from the literature relating market concentration to intangible investments.[9]

2. Empirical facts

In this section, we document the firm-level relationships between innovation, advertising, and market share for large firms in the United States. To do so, we obtain financial statement data on public firms from Compustat, and link them to patent micro-data from the US Patent and Trademark Office (USPTO). These relationships motivate our theoretical framework which, unlike earlier papers in the literature, can both generate and quantitatively match the new stylized facts that we uncover in this section. For brevity, we relegate the details on data sets and variable construction to Appendix A.

Theoretically, the relationship between competition and innovation is ambiguous. Intense competition from peers can encourage a firm to innovate to escape competition, which would imply a positive relationship. At the same time, competition pushes down profitability, which discourages innovation — the so-called Schumpeterian effect. Similar forces are at work when one considers the relationship between competition and advertising: competition from peers can trigger a rat race in advertising. But at the same time, if profitability is low, returns to advertising are likewise lower.

Our empirical analysis reveals that both relationships are non-monotonic in the data. As a first pass, we plot the raw relationships in Fig. 1. The left panel demonstrates an inverted-U relationship between a firm's innovation (as captured by the average citations of the patents it applies for in a given year) and the firm's sales relative to all other US listed firms in its 4-digit SIC industry. The middle and right panels are similar figures for log R&D spending and log advertising, respectively.

For a more rigorous assessment, we run the regressions

$$AvgCite_{ijt} = \text{const.} + \beta_1^{inn}\sigma_{ijt} + \beta_2^{inn}\sigma_{ijt}^2 + \boldsymbol{X}'_{ijt}\boldsymbol{\gamma}^{inn} + \alpha_j + \alpha_t + u_{ijt} \tag{1a}$$

$$LogAdvExp_{ijt} = \text{const.} + \beta_1^{adv}\sigma_{ijt} + \beta_2^{adv}\sigma_{ijt}^2 + \boldsymbol{X}'_{ijt}\boldsymbol{\gamma}^{adv} + \alpha_j + \alpha_t + u_{ijt} \tag{1b}$$

for firm i in 4-digit SIC industry j and year t, where the outcome variables are either average citations (our baseline measure of innovation) or log advertising expenditures (constructed using the xad variable in Compustat).[10] On the right-hand side, σ_{ijt} denotes firm i's relative sales within its industry, σ_{ijt}^2 is market share squared, (α_j, α_t) are industry and year fixed effects, and $\boldsymbol{X}_{ijt}$ is a set of firm-level controls, including profitability, leverage, market-to-book ratio, log R&D stock, firm age, the coefficient of variation of the firm's stock price, and the number of firms in the industry.[11] For every one of our specifications, we normalize the dependent variables by subtracting their means and dividing by their standard deviation. The regression results are reported in Table A.1. As

9 The features of our model link our paper to the broader macroeconomics literature on competition and markups (see e.g., Covarrubias et al. (2020), Gutiérrez et al. (2021)).

10 The results for regression (1a) are robust to alternative innovation measures such as patent count, patent quality, tail innovations, originality, generality, R&D expenses, as well as investments that are potentially correlated with innovation (physical capital investment) and direct measures of firm growth (sales growth, employment growth, asset growth). The results for regression (1b) are robust to using the logarithm of Selling, General, and Administrative Expenses (SG&A) as a proxy for advertising.

11 To alleviate concerns regarding the effects of a persistent customer base accumulation on this relationship, in addition to adding firm age as a control, we also create a persistent demand stock variable following Fitzgerald et al. (2023), and add it as an additional linear control. Our results are robust to its inclusion, which is shown in Table F.3 in the Online Appendix.

can be seen, the coefficients of the linear terms are positive, and those of the square terms are negative, indicating an inverted-U relationship once again.[12]

Finally, we are also interested in how the relative sales of a firm are associated with the firm's own advertising efforts versus those of its competitors. Table A.2 reports results for the regression:

$$\sigma_{ijt} = \text{const.} + \beta_3 \; LogAdvExp_{ijt} + \beta_4 \; OthersLogAdv_{ijt} + \boldsymbol{X}'_{ijt}\boldsymbol{\delta} + \alpha_j + \alpha_t + \varepsilon_{ijt} \tag{2}$$

where $OthersLogAdv_{ijt}$ is a measure of the total advertising spending in the industry, excluding firm i. The first column of Table A.2 reports results when $OthersLogAdv_{ijt}$ is computed as aggregate advertising expenditures by other firms in the same industry. The second column does the same using the average advertising of the competitors instead of the sum. We use the same controls and the same normalization for advertising variables as in Table A.1. Both measures of competitor advertising convey the same picture: a firm's own advertising is positively correlated with its relative sales, whereas that of competitors is negatively correlated. The coefficients are roughly comparable in magnitude.

Inspired by these findings, in the next section we develop an oligopolistic general-equilibrium growth model with firm heterogeneity in which innovation and advertising decisions strategically interact.

3. Model

3.1. Environment

Preferences. Time is continuous, and indexed by $t \in \mathbb{R}_+$. The economy is populated by an infinitely-lived representative consumer who maximizes lifetime utility:

$$W = \int_0^{+\infty} e^{-\rho t} \ln(C_t) \; dt \tag{3}$$

where $\rho > 0$ is the time discount rate, and C_t is consumption of final good at time t. The price of the final good is normalized to one. The household is endowed with one unit of time every instant, supplied inelastically to the producers in the economy in return for a wage w_t, which clears the labor market. The household owns all the firms in the economy and carries a stock of wealth $\mathcal{A}_t$ each period, equal to the total value of corporate assets. The budget constraint satisfies $\dot{\mathcal{A}}_t = r_t\mathcal{A}_t + w_t - C_t$, where r_t is the rate of return on assets.

Final good production. The final good Y_t is produced by a competitive representative firm using inputs from a measure one of industries, with technology:

$$Y_t = \exp\left(\int_0^1 \ln\left(y_{jt}\right) \; dj\right) \tag{4}$$

where y_{jt} is production of industry j at time t.

Industry production. Each industry j is populated by an endogenous number of superstar firms, $N_{jt} \in \{1, \dots, \bar{N}\}$, each producing a differentiated variety, as well as by a competitive fringe composed of an endogenous mass m_{jt} of small firms producing a homogeneous good. Industry j's output at time t is given by:

$$y_{jt} = \left(\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}} + \tilde{y}_{sjt}^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}} \tag{5}$$

where $\tilde{y}_{cjt}$ denotes the output of the fringe, $\tilde{y}_{sjt}$ denotes the output of superstars, and $\gamma > 1$ is the elasticity of substitution between the two. Fringe firms produce a homogeneous product, so $\tilde{y}_{cjt} = \int_{F_{jt}} y_{ckjt} \; dk$, where F_{jt} is the endogenous set of small firms in the fringe in industry j at time t.[13] Given that there is a continuum of small firms and their products are homogeneous, each small firm in the competitive fringe is a price-taker.

By contrast, superstar firms behave strategically, competing in quantities in a static Cournot game, as in Atkeson and Burstein (2008). Total production by superstars of industry j at time t is given by:

$$\tilde{y}_{sjt} = \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} y_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \tag{6}$$

[12] To further establish the robustness of our results, we conduct the hypothesis test proposed in Lind and Mehlum (2010) for regressions in Table A.1, where the null hypothesis is the lack of an inverted-U relationship. This involves testing whether or not the slope of the curve is positive at the start and negative at the end of the interval of the variable of interest. Correspondingly, Table A.3 reports the t- and p-values at the lower and upper bounds of the interval of the explanatory variable. The null hypothesis is firmly rejected in both specifications. The inverted-U relationships that we have identified pass the formal test of existence, with p-values below 1% in both regressions.

[13] The competitive fringe allows us to have a realistic firm size distribution in each industry. Industries in the US are populated by thousands of firms on average. While a handful of superstar firms account for a large fraction of the total industry sales, collectively, the remaining small firms also account for a significant share even if their individual market shares are minuscule. Our calibrated model matches this (see Section 4.2).

where $\eta > 1$ is the elasticity of substitution between varieties, holding $\eta > \gamma$. Each variety has perceived quality $\hat{\omega}_{ijt}$, defined by:

$$\hat{\omega}_{ijt} = \frac{1 + \omega_{ijt}}{\frac{1}{N_{jt}} \sum_{k=1}^{N_{jt}} (1 + \omega_{kjt})} \tag{7}$$

In this expression, ω_{ijt} is a quality shifter which is affected by the superstar firm's advertising decisions, as described below. The perceived quality of a product, $\hat{\omega}_{ijt}$, is the ratio of this quality shifter to the average shifter among the superstars within the industry. Intuitively, we model advertising as a technology that allows firms to shift the perceived quality of their own product.[14] Moreover, all else equal, if a firm chooses to increase its advertising efforts, it will increase the perceived quality of its own product while decreasing that of every other product in the same industry. In this sense, advertising is akin to a zero-sum game, in which a firm's advertising efforts are directly detrimental to other firms' product qualities, so that if all superstars were to choose any identical ω level, then all varieties would have the same baseline perceived quality.[15]

Firms' production technology. In each industry, superstar firms and small firms in the fringe produce their variety using a linear production technology in labor: $y_{ijt} = q_{ijt} l_{ijt}$ and $y_{ckjt} = q_{cjt} l_{ckjt}$, respectively, where q_{ijt} is the productivity of superstar firm i in industry j at time t, and q_{cjt} is the productivity of a fringe firm. Each small firm from the fringe is assumed to have the same productivity within an industry. Superstar firms, by contrast, are heterogeneous in their productivity, which can be built over time through innovation.

R&D and innovation. Each superstar can perform R&D to improve the productivity of its variety. To generate a Poisson rate z_{ijt} of success in R&D, firm i must pay $R_{ijt} = \chi z_{ijt}^{\phi} Y_t$ units of the final good, where $\chi > 0$ and $\phi > 1$ are parameters. A successful innovator is able to advance its productivity by a factor $(1+\lambda)$, where $\lambda > 0$. As we shall see shortly, industry-level outcomes in this model depend on the *relative* levels of productivity between superstar firms, which can be summarized by an integer $n_{ijt}^k \in \{-\bar{n}, -\bar{n}+1, \dots, \bar{n}-1, \bar{n}\}$ holding $q_{ijt}/q_{kjt} = (1+\lambda)^{n_{ijt}^k}$. In words, n_{ijt}^k is the number of productivity steps by which firm i in industry j is ahead (if $n_{ijt}^k > 0$), behind (if $n_{ijt}^k < 0$), or neck-to-neck (if $n_{ijt}^k = 0$) with respect to firm $k \neq i$ at time t. The parameter $\bar{n} \geq 1$ is the maximum number of steps between any two superstar firms within an industry. For the competitive fringe, we assume that the relative productivity of small firms with respect to the leader is a constant, denoted by the parameter $\zeta = \frac{q_{cjt}}{q_{jt}^{leader}}$, where $q_{jt}^{leader} \equiv \max_{k=1,\dots,N_{jt}} \{q_{kjt}\}$.

Advertising. Each superstar firm can spend resources on advertising its product to affect perceived quality $\hat{\omega}_{ijt}$. In order to achieve a quality shifter ω_{ijt}, firm i of industry j must spend $A_{ijt} = \chi_a \omega_{ijt}^{\phi_a} Y_t$ units of the final good, where $\chi_a > 0$ and $\phi_a > 1$ are parameters.

Entry and exit of superstar firms. At any time t, each small firm k in the competitive fringe can generate a Poisson arrival density X_{kjt} and enter into the pool of superstar firms, as long as $N_{jt} < \bar{N}$ for some $\bar{N}$ set exogenously. The associated R&D cost, expressed in units of the final good, is given by $R_{kjt}^e = \nu X_{kjt}^{\epsilon} Y_t$, with $\nu > 0$ and $\epsilon > 1$. As small firms are all homogeneous within the same industry, their level of innovation is identical in equilibrium. This allows us to write an industry-level Poisson rate of innovation as $X_{jt} = \int X_{kjt} dk = m_{jt} X_{kjt}$. Similarly, the R&D expenditures of small firms at the industry level equal $R_{jt}^e = m_{jt} R_{kjt}^e$.

Upon successful entry (provided $N_{jt} < \bar{N}$), the entrant is assumed to enter as the smallest superstar firm within the industry and thus becomes a superstar firm with productivity level $\bar{n}$ steps behind the leader. In this case, the number of superstar firms N_{jt} increases by one. On the other hand, a superstar firm endogenously loses its superstar status when it falls more than $\bar{n}$ steps below the industry leader. In that case, N_{jt} decreases by one.

Entry and exit of small firms. Finally, there is entry into, and exit out of, the competitive fringe. We assume an exogenous exit rate of small firms equal to $\tau > 0$. For entry, we assume that there is a measure one of entrepreneurs who pay $\psi e_t^2 Y_t$ units of the final good to generate a Poisson rate e_t of starting a new small firm, where $\psi > 0$. New firms are randomly allocated to the competitive fringe of an industry, implying $m_{jt} = m_t$ for all industries j so long as $m_{j0} = m_0$. We further assume that successful entrepreneurs sell their firm on a competitive market at its full value and remain in the set of entrepreneurs, which keeps the mass of entrepreneurs unchanged.

3.2. Equilibrium

Household's problem. Utility maximization delivers the Euler equation $\frac{\dot{C}_t}{C_t} = r_t - \rho$.

---

[14] We should note that we model advertising as a static decision. The marketing literature typically assumes that advertising has a carry-over effect, i.e., current advertising affects future sales through an advertising stock of goodwill. However, this literature also shows that this effect is short-lived, and that it depreciates within weeks, becoming virtually zero in a year. See among others Leone (1995), Dubé et al. (2005), Doganoglu and Klapper (2006), Danaher et al. (2008), Terui et al. (2011), Shapiro et al. (2021), and Bagwell (2007) for an overview of the literature. Our model's focus is on long-term effects rather than high-frequency variations shorter than a year. In addition, advertising through goodwill would still act as a demand shifter and lead to mechanisms similar to those obtained in our current model. This is not to say that there are no mechanisms other than advertising that can generate persistence in the demand for a firm's products (i.e., a persistent customer base). See, for instance, Gourio and Rudanko (2014), Afrouzi et al. (2023), Roldan-Blanco and Gilbukh (2021), or Ignaszak and Sedláček (2022).

[15] In Section 6, we relax this assumption and allow for advertising to be non-combative in an extended model, removing the zero-sum game property. All our main findings are robust to this change, although the exact magnitudes change.

Final good producers. The final good is produced competitively. The representative final good producer chooses the quantity of each variety in each industry to achieve a given level of output which minimizes total production costs. This leads to the following demand functions for superstar and fringe firms, respectively[16]:

$$y_{ijt} = \hat{\omega}_{ijt}^{\eta}\left(\frac{p_{ijt}}{\tilde{p}_{sjt}}\right)^{-\eta}\left(\frac{\tilde{p}_{sjt}}{p_{jt}}\right)^{-\gamma}\frac{Y_t}{p_{jt}} \tag{8a}$$

$$\tilde{y}_{cjt} = \left(\frac{\tilde{p}_{cjt}}{p_{jt}}\right)^{-\gamma}\frac{Y_t}{p_{jt}} \tag{8b}$$

where p_{ijt} is the price of the variety produced by superstar i in industry j at time t, and $\tilde{p}_{cjt}$ is the price of the homogeneous product of the competitive fringe of that industry. Additionally, we have defined $\tilde{p}_{sjt} \equiv [\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt}^{\eta} p_{ijt}^{1-\eta}]^{\frac{1}{1-\eta}}$ as the ideal price index among the different varieties of the superstars, and $p_{jt} \equiv [\tilde{p}_{sjt}^{1-\gamma} + \tilde{p}_{cjt}^{1-\gamma}]^{\frac{1}{1-\gamma}}$ as the ideal price index of the industry. The relative output between two superstars i and k of the same industry is:

$$\frac{y_{ijt}}{y_{kjt}} = \left(\frac{\hat{\omega}_{kjt}}{\hat{\omega}_{ijt}}\frac{p_{ijt}}{p_{kjt}}\right)^{-\eta} \tag{9}$$

This makes it apparent that firms can use advertising to shift demand towards their products and thereby increase profits at the expense of their direct competitors. The allocation of expenditure between superstars and small firms within the same industry is determined by the relative price index $\frac{\tilde{p}_{sjt}}{p_{jt}}$, with price-elasticity γ. In particular, the relative output between a superstar and the industry's fringe is:

$$\frac{y_{ijt}}{\tilde{y}_{cjt}} = \hat{\omega}_{ijt}^{\eta}\left(\frac{p_{ijt}}{\tilde{p}_{sjt}}\right)^{-\eta}\left(\frac{\tilde{p}_{sjt}}{\tilde{p}_{cjt}}\right)^{-\gamma} \tag{10}$$

Market shares and markups. Each superstar firm simultaneously chooses output (y_{ijt}) and advertising (ω_{ijt}) to maximize profit:

$$\max_{y_{ijt},\omega_{ijt}} \left\{p_{ijt}y_{ijt} - w_t l_{ijt} - \chi_a \omega_{ijt}^{\phi_a} Y_t\right\} \tag{11}$$

subject to Eqs. (8a)–(8b), and $y_{ijt} = q_{ijt} l_{ijt}$. As superstar firms within the same industry compete *à la* Cournot, they internalize how their output choices affect the aggregate output within their industry. In equilibrium, each superstar firm i sets a markup over the marginal cost, so the price equals $p_{ijt} = M_{ijt}\frac{w_t}{q_{ijt}}$, and the equilibrium markup is given by:

$$M_{ijt} = \left[\left(\frac{\eta-1}{\eta}\right) - \left(\frac{\gamma-1}{\gamma}\right)\sigma_{ijt} - \left(\frac{\eta-\gamma}{\eta\gamma}\right)\tilde{\sigma}_{ijt}\right]^{-1} \tag{12}$$

In this formula, we have defined:

$$\sigma_{ijt} \equiv \frac{p_{ijt}y_{ijt}}{p_{jt}y_{jt}} \quad \text{and} \quad \tilde{\sigma}_{ijt} \equiv \frac{p_{ijt}y_{ijt}}{\tilde{p}_{sjt}\tilde{y}_{sjt}} \tag{13}$$

as, respectively, the market share of firm i among all firms (including superstars and fringe) in its industry, and the market share of the firm among the superstars only. Eq. (12) shows that a firm's markup is increasing in both of these market share definitions.

We can write market shares in terms of relative outputs and productivities, which allows us to obtain the following set of static equilibrium conditions:

$$\left(\frac{y_{ijt}}{y_{kjt}}\right)^{\frac{1}{\eta}} = \frac{q_{ijt}}{q_{kjt}}\frac{\hat{\omega}_{ijt}}{\hat{\omega}_{kjt}}\frac{M_{kjt}}{M_{ijt}}, \ \forall k \neq i \tag{14a}$$

$$\frac{y_{ijt}}{\tilde{y}_{cjt}} = \frac{q_{ijt}}{q_{cjt}}\frac{\sigma_{ijt}}{\sigma_{cjt}}\frac{1}{M_{ijt}}. \tag{14b}$$

where $\sigma_{cjt} \equiv 1 - \sum_{k=1}^{N_{jt}} \sigma_{kjt}$ is the market share of the fringe. In words, the relative demand of superstars is increasing in their relative productivity and relative taste shifter, and decreasing in their relative markup. Static profits before advertising costs ($\pi_{ijt} \equiv p_{ijt}y_{ijt} - w_t l_{ijt}$) are proportional to the superstar's market share and the Lerner index, $\pi_{ijt} = \sigma_{ijt}(1 - M_{ijt}^{-1})Y_t$.

Advertising choices. As with output choices, a superstar firm internalizes that its advertising decisions affect industry-level prices through their impact on the firm's own market shares relative to other superstars and the fringe, as well as on other firms' perceived quality ($\hat{\omega}_{kjt}$). The optimal level of advertising ω_{ijt} by firm i in industry j equates the marginal static profit gains from advertising to the marginal cost of advertising. Deriving the first-order condition with respect to ω_{ijt} (see Appendix C.2 for details), we can write:

$$\frac{\sigma_{ijt}}{1+\omega_{ijt}}\left[\frac{N_{jt}-\hat{\omega}_{ijt}}{N_{jt}} + \frac{\gamma-\eta}{(\eta-1)\gamma}\left(\tilde{\sigma}_{ijt} - \frac{\hat{\omega}_{ijt}}{N_{jt}}\right) + \frac{\eta}{\eta-1}\frac{\gamma-1}{\gamma}\sigma_{ijt}\left(\frac{\hat{\omega}_{ijt}}{N_{jt}\tilde{\sigma}_{ijt}} - 1\right)\right] = \chi_a\phi_a\omega_{ijt}^{\phi_a-1} \tag{15}$$

[16] See Appendix C.2 for a detailed derivation of the set of all static equilibrium conditions.

As both markups and taste shifters are functions of market shares, and these are themselves functions of relative outputs, Eqs. (14a)–(14b) and (15) comprise a system of $2N_{jt}$ equations in $2N_{jt}$ unknowns (the output ratios and advertising decisions), which can be solved, for each industry j, as a function of the set of relative productivities between firms, $\{n^k_{ijt}\}$, and the total number of superstars in the industry, N_{jt}.

We denote equilibrium post-advertising profits by $\pi^{adv}_{ijt} \equiv \pi_{ijt} - \chi_a \omega^{\phi_a}_{ijt} Y_t$, which will drive the incentives for firms to invest in R&D and innovation.

Labor market clearing. We close the static part of the equilibrium by imposing labor market clearing. Labor input choices satisfy $l_{ijt} = \frac{\sigma_{ijt}}{w^{rel}_t} M^{-1}_{ijt}$ and $l_{cjt} = \frac{\sigma_{cjt}}{w^{rel}_t}$, for each superstar firm i and the fringe, respectively, where $w^{rel}_t \equiv \frac{w_t}{Y_t}$ denotes the relative wage. Imposing labor market clearing gives us a formula for this relative wage:

$$w^{rel}_t = \int_0^1 \left(\sigma_{cjt} + \sum_{i=1}^{N_{jt}} \sigma_{ijt} M^{-1}_{ijt} \right) dj \tag{16}$$

Superstar value function and R&D decision. The relevant state for firm i in industry j is given by the vector collecting the number of productivity steps relative to all other superstars in the industry, $\mathbf{n}_{ijt} \equiv \{n^k_{ijt}\}_{k\neq i}$, and the number of superstars in the industry, $N_{jt} \equiv |\mathbf{n}_{ijt}| + 1$. We drop time and industry subscripts unless otherwise needed. A superstar firm i chooses an innovation rate z_i to maximize the value of the firm, given by:

$$\begin{aligned} rV(\mathbf{n}_i, N) = \max_{z_i} \Bigg\{ & \pi^{adv}(\mathbf{n}_i, N) - \chi z_i^{\phi} Y \\ & + z_i \left[V\left(\mathbf{n}_i \backslash \{n^k_i = \bar{n}\} + \mathbf{1}, N - |\{n^k_i = \bar{n}\}|\right) - V(\mathbf{n}_i, N) \right] - \sum_{\{k: n^k_i = -\bar{n}\}} z_{kj} V(\mathbf{n}_i, N) \\ & + \sum_{\{k: n^k_i > -\bar{n}\}} z_{kj} \left[V\left(\mathbf{n}_i \backslash \{n^k_i\} \cup \{n^k_i - 1\} \backslash \{n^l_i = \bar{n} + n^k_i\}, N - |\{n^l_i = \bar{n} + n^k_i\}|\right) - V(\mathbf{n}_i, N) \right] \\ & + X_j \left[V\left(\mathbf{n}_i \cup \left\{\min\left(\bar{n}, \bar{n} + \min(\mathbf{n}_i)\right)\right\}, \min(N+1, \bar{N})\right) - V(\mathbf{n}_i, N) \right] \Bigg\} + \dot{V}(\mathbf{n}_i, N) \end{aligned} \tag{17}$$

In this Hamilton–Jacobi-Bellman equation, the first line is the profit flow from sales net of labor and advertising costs, minus the costs of R&D. The first term on the second line is the gain from a successful innovation, which increases the lead of the firm by one step relative to all of its competitors. Any firm $\bar{n}$ productivity steps below firm i exits the set of superstars, which decreases the number of superstars by one. The second term on this line is the change in value due to endogenously exiting the set of superstars after a successful innovation by an industry leader who is $\bar{n}$ steps ahead of firm i, in case any such firm exists. The third line includes the event that any other superstar k of the industry innovates. In this case, the lead of firm i relative to the innovating firm decreases by one. Moreover, in case the innovating firm was leading any other firm l by $\bar{n}$, such firm exits, and the number of superstars in the industry decreases. The first term on the fourth line is the effect of the emergence of a new superstar on the value of firm i, with the incoming firm starting with distance $\bar{n}$ from the industry leader. The last term on this line is the change in firm value over time.

In a balanced growth path (BGP) with constant output growth $g > 0$, we have $V(\mathbf{n}_i, N) = v(\mathbf{n}_i, N)Y$ for a time-invariant function v. Using the Euler equation, we can write:

$$\begin{aligned} \rho v(\mathbf{n}_i, N) = \max_{z_i} \Bigg\{ & \frac{\pi^{adv}(\mathbf{n}_i, N)}{Y} - \chi z_i^{\phi} + z_i \left[v\left(\mathbf{n}_i \backslash \{n^k_i = \bar{n}\} + \mathbf{1}, N - |\{n^k_i = \bar{n}\}|\right) - v(\mathbf{n}_i, N) \right] \\ & + \sum_{\{k: n^k_i \neq -\bar{n}\}} z_{kj} \left[v\left(\mathbf{n}_i \backslash \{n^k_i\} \cup \{n^k_i - 1\} \backslash \{n^l_i = \bar{n} + n^k_i\}, N - |\{n^l_i = \bar{n} + n^k_i\}|\right) - v(\mathbf{n}_i, N) \right] \\ & - \sum_{\{k: n^k_i = -\bar{n}\}} z_{kj} v(\mathbf{n}_i, N) + X_j \left[v(\mathbf{n}_i \cup \{\min\left\{\bar{n}, \bar{n} + \min(\mathbf{n}_i)\right\}\}, \min(N+1, \bar{N})) - v(\mathbf{n}_i, N) \right] \Bigg\} \end{aligned} \tag{18}$$

Optimal innovation is $z_i = \left[(v(\mathbf{n}_i \backslash \{n^k_i = \bar{n}\} + \mathbf{1}, N - |\{n^k_i = \bar{n}\}|) - v(\mathbf{n}_i, N)) / (\chi \phi) \right]^{\frac{1}{\phi - 1}}$.

Small firm innovation. To obtain the optimal behavior of small firms and entry into the superstar status, we define $\Theta \equiv (N, \vec{n})$ as the state of an industry, where $N \in \{1, \dots, \bar{N}\}$ is the number of superstars in the industry and $\vec{n} \in \{0, \dots, \bar{n}\}^{N-1}$ is the number of steps followers are behind the leader (in ascending order). Further, define $p(\Theta, \Theta')$ as the instantaneous flows from state Θ to Θ'. In each industry j of type Θ (with $N(\Theta) < \bar{N}$ superstars), each small firm in the competitive fringe chooses R&D investment to maximize:

$$\begin{aligned} rV^e(\Theta) = \max_{X_{kj}} \Bigg\{ & X_{kj} V\left(\{\tilde{\mathbf{n}}_j - \bar{n}\} \cup \{-\bar{n}\}, N_j + 1\right) - \tau V^e(\Theta) - \nu X^{\epsilon}_{kj} Y \\ & + \sum_{\Theta'} p(\Theta, \Theta') \left(V^e(\Theta') - V^e(\Theta) \right) \Bigg\} + \dot{V}^e(\Theta) \end{aligned} \tag{19}$$

where $V^e(\Theta)$ is the value of a small firm in industry j and $\tilde{\mathbf{n}}_j$ is the firm state variable of the industry leader. Guessing and verifying that $V^e(\Theta) = v^e(\Theta)Y$ in a BGP, the optimal innovation intensity by a small firm in industry j is then $X_{kj} = \left[v(\{\tilde{\mathbf{n}}_j - \bar{n}\} \cup \{-\bar{n}\}, N_j + 1)/(\nu\epsilon)\right]^{\frac{1}{\epsilon-1}}$.

Plugging in the optimal solution, the normalized value of a small firm is:

$$v^e(\Theta) = \frac{1}{\rho+\tau}\left[\left(1-\frac{1}{\epsilon}\right)\frac{v(\{\tilde{\mathbf{n}}_j - \bar{n}\} \cup \{-\bar{n}\}, N_j+1)^{\frac{\epsilon}{\epsilon-1}}}{(\nu\epsilon)^{\frac{1}{\epsilon-1}}} + \sum_{\Theta'} p(\Theta,\Theta')\Big(v^e(\Theta') - v^e(\Theta)\Big)\right] \tag{20}$$

Entrepreneurs. The expected value of a new small firm created by a successful entrepreneur is equal to $W = \sum_\Theta V^e(\Theta)\mu(\Theta)$, where $\mu(\Theta)$ is the equilibrium measure of industries of type Θ. The value of being an entrepreneur, denoted S, is:

$$rS = \max_e\{-\psi e^2 Y + eW\} + \dot{S} \tag{21}$$

Guessing and verifying that $S = sY$ in a BGP, we obtain that $e = \frac{1}{2\psi}\sum_\Theta v^e(\Theta)\mu(\Theta)$, which implies $s = \frac{1}{4\psi\rho}\left[\sum_\Theta v^e(\Theta)\mu(\Theta)\right]^2$. In a BGP, entry into the competitive fringe equals exit from the competitive fringe, implying $e = \tau m$. Thus, we get the equilibrium measure of small firms in the economy:

$$m = \frac{\sum_\Theta v^e(\Theta)\mu(\Theta)}{2\psi\tau} \tag{22}$$

We focus on the unique Markov Perfect Equilibrium of this economy, which is formally described in Appendix C.1.

Growth rate. The BGP rate of economic growth is given by[17]:

$$g = \ln(1+\lambda)\sum_\Theta p^{leader}(\Theta)\mu(\Theta) \tag{23}$$

where $p^{leader}(\Theta)$ is defined as the arrival rate of leader innovation. In words, the growth rate of the economy is the product of the log step size of innovations and the average leader innovation intensity across industries. Innovation by all other firms affects this growth rate through its influence on the stationary industry state distribution $\mu(\Theta)$ and its strategic effect on leader innovation.

4. Calibration strategy

Our main goal in this paper is to understand the static and dynamic welfare implications of advertising through its interaction with innovation and the market share distribution within and across industries. To quantify these effects, our calibration strategy requires that the model is consistent with empirical observations regarding advertising, innovation, and market concentration. In particular, we make sure that the model is able to replicate the micro-level empirical regularities documented in Section 2, as well as other salient properties of the aggregate data.

4.1. Indirect inference and model fit

We calibrate the model at the annual frequency, and set the consumer discount rate externally to $\rho = 0.04$. This leaves 12 parameters to calibrate: $(\lambda, \eta, \gamma, \chi, \nu, \zeta, \phi, \epsilon, \tau, \psi, \chi_a, \phi_a)$.[18] Individual moments from the data cannot uniquely identify each parameter separately. We therefore calibrate the 12 parameters jointly through a moment-matching procedure, by ensuring that the distance between the model-generated moments and their empirical counterparts is minimized. The identification success of this method requires that we choose moments which are sufficiently sensitive to variation in the structural parameters. In what follows, we offer some verbal intuition for this identification. Appendix B.3 includes a full discussion, as well as a rigorous global identification test.

We target a combination of aggregate, industry-level, and firm-level moments. Appendix B.2 provides details regarding the data sources and the way these moments are computed both in the US data and in the model. At the aggregate level, we target the growth rate of real GDP per capita, which helps us identify the step size of innovations, λ. To make sure the model correctly predicts the degree of market power in the economy, we also target the sales-weighted average and standard deviation of firm-level markups, the average firm profitability, and the aggregate labor share. These moments are informative about the within-industry elasticity of substitution between superstars and fringe, γ, and the relative productivity of the fringe, ζ. The R&D cost parameters for small and large firms are identified using aggregate and industry-level moments related to leader quality and R&D expenditures. The average relative quality of industry leaders, and its standard deviation across industries, help us pin down the scale parameter for the small firms' R&D cost function, ν, and the curvature in the R&D costs of superstars, ϕ.[19] The scale parameter in the R&D cost function of superstars, χ, is pinned down by the aggregate R&D share of GDP. Similarly, the scale parameter in the advertising cost function,

[17] See Appendix C.3 for the derivation.

[18] To calibrate the model, we must also choose values for the maximum number of steps between superstars, $\bar{n}$, and the maximum number of superstars in an industry, $\bar{N}$. We set $\bar{n} = 5$ and $\bar{N} = 4$, which imply that the dimensionality of the firm state variable is 774, whereas that of the industry state variable is 84. In Section 6 we perform robustness checks for both $\bar{n}$ and $\bar{N}$.

[19] We proxy quality with the stock of past patent citations, and define the relative quality of the leader as the quality of the top firm divided by the sum of the qualities of the top four firms in its SIC4 industry. See Appendix B.2 for details.

Table 1
Benchmark model parameters and target moments.

A. Calibrated parameters

Parameter	*Description*	*Value*
λ	Innovation step size	0.1657
η	Elasticity within industry	11.6743
γ	Elasticity between superstars and fringe	2.9637
χ	Superstar cost scale	77.4786
ν	Small firm cost scale	3.1629
ζ	Competitive fringe ratio	0.7078
ϕ	Superstar cost convexity	4.4849
ϵ	Small firm cost convexity	4.5514
τ	Small firm exit rate	0.1151
ψ	Entry cost scale	0.0597
χ_a	Advertising cost scale	0.0664
ϕ_a	Advertising cost convexity	3.3646

B. Moments

Target moments	*Data*	*Model*
Growth rate	2.204%	2.201%
R&D/GDP	2.435%	2.467%
Advertising/GDP	2.200%	2.208%
Average markup	1.350	1.342
Standard deviation of markups	0.346	0.442
Labor share	0.652	0.638
Firm entry rate	0.115	0.115
Average profitability	0.144	0.136
Average leader relative quality	0.749	0.510
Standard deviation of leader relative quality	0.223	0.164
Regression (1a), linear coefficient: β_1^{inn}	0.629	0.982
Regression (1a), top point: $-\beta_1^{inn}/(2\beta_2^{inn})$	0.505	0.483
Regression (1b), linear coefficient: β_1^{adv}	6.260	7.614
Regression (1b), top point: $-\beta_1^{adv}/(2\beta_2^{adv})$	0.533	0.521

Notes: Panel A reports the calibrated parameters. Panel B reports the simulated and empirical moments. Appendix B.2 provides information on our data and the way these moments are computed in the data and the model. Appendix B.3 discusses identification.

χ_a, is pinned down by the aggregate advertising share. The exit rate of small firms, τ, is directly identified by targeting the firm entry rate in the data.

The remaining moments ensure that the model reproduces the facts documented in Section 2, namely the two non-linear relationships between innovation and market share, and between advertising and market share. To discipline these intra-industry inverted-U shaped relationships, we target (i) the linear coefficients of both regressions (1a) and (1b), and (ii) the top points of the inverted U, which we can obtain using a combination of the linear and quadratic coefficients, $-\beta_1^{inn}/(2\beta_2^{inn})$ and $-\beta_1^{adv}/(2\beta_2^{adv})$, respectively. These coefficients help pin down the elasticity of substitution between superstars, η, the curvature in the R&D cost function of small firms, ϵ, and the curvature in the advertising cost function, ϕ_a, as we explain in Appendix B.3. Finally, given this assignment, the value for ψ is found ex-post to normalize the stationary measure of small firms, m, to one in equilibrium.

Table 1 presents the results of our calibration exercise. Panel A reports the parameter values, and Panel B reports the results in terms of moment matching. The model manages to match the targeted moments well, even though the model is over-identified.

4.2. Model validation and untargeted moments

Besides matching the targeted moments, our calibrated model is also consistent with the data along several untargeted dimensions that are relevant for our analysis. First, our calibration generates concentration ratios that are in line with the data. The average combined market share of the largest four firms within their industry (CR4 ratio) is 43.13% in our model, similar to its empirical counterpart (48.41%). The overall distribution of CR4 ratios across industries is also closely matched, as seen in Table F.2.

Second, the model aligns well with the observed empirical firm-level correlations between advertising, R&D, SG&A (i.e., Selling, General, and Administrative Expenses), markups, and profitability. De Loecker et al. (2020) report that both firm-level markups and profitability are positively correlated with R&D, advertising, and SG&A, which our model also delivers (see Table F.1).[20]

Third, in the data we find that the advertising of competitors within the industry is negatively associated with a firm's market share, i.e., $\beta_4 < 0$ in regression (2). In the model, demand for a firm's variety increases with its advertising effort, but it is also

[20] Our model delivers correlations that are slightly higher than those reported in De Loecker et al. (2020), a result that can be attributed to the presence of noise in the data.

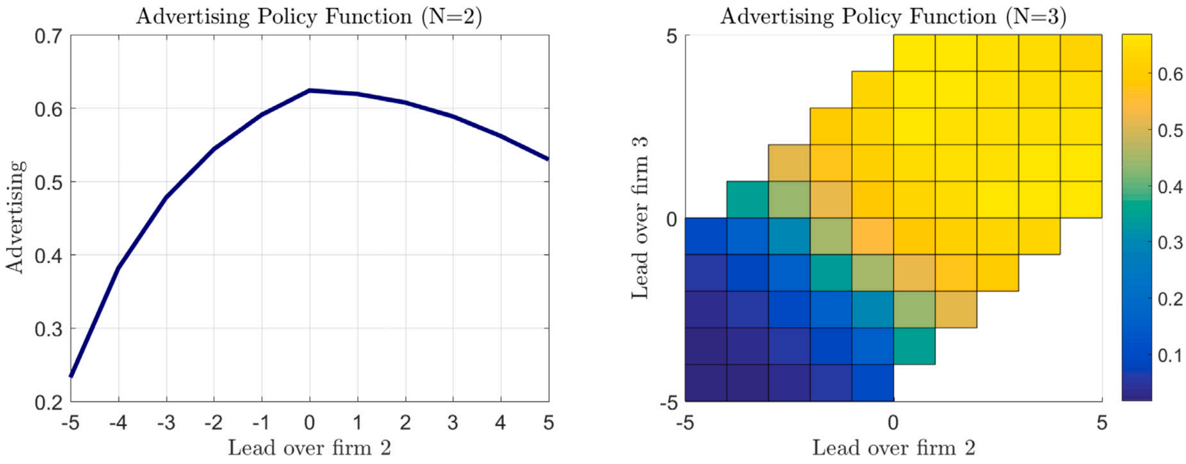


Fig. 2. Advertising policy function.
Notes: This figure displays the policy functions for advertising as functions of this firm's technological lead relative to its competitor(s) for the case of industries with $N = 2$ superstar firms (left panel) and $N = 3$ superstar firms (right panel).

negatively affected by the advertising spending of the competitors. Regressing a firm's relative sales on its own advertising and the total advertising of its competitors in the model allows us to pin down the relative role played by both types of advertising on sales. Measuring this relative role by the ratio of the two coefficients, β_3/β_4, our calibrated model delivers a value of -1.28, which is comparable with the one observed in the data, -1.01.

Fourth, heterogeneous markups arise in our framework, since the price-elasticity of demand is size-dependent. Larger firms face more inelastic demand and are able to charge higher markups. As a result, the relationship between demand elasticity and firm size (the so-called *superelasticity* of demand) shapes the distribution of markups along relative firm size and governs the response of markups to policy in our counterfactual experiments. The average superelasticity in our calibrated model equals 0.21, which is very similar to the superelasticity of 0.16 estimated by Edmond et al. (2023).[21]

Finally, our calibrated model is also consistent with the data regarding the correlation between labor shares and value added. Gouin-Bonenfant (2022) documents a negative relationship between (log) firm-level labor share and (log) value-added. Repeating this regression in our calibrated model, we obtain a coefficient of −0.0648, which is close to the value −0.112 that Gouin-Bonenfant (2022) estimates for Canada. Gouin-Bonenfant (2022) also documents a negative correlation between the industry-level labor share and the dispersion of productivity, which our calibrated model also delivers.

4.3. Key model features

Before proceeding to our quantitative analysis, this section briefly discusses the key features of the equilibrium that arise from our calibrated model.

Optimal advertising and innovation policies. Fig. 2 presents the policy functions for advertising for the case of industries with $N = 2$ superstar firms (left panel) and $N = 3$ superstar firms (right panel).[22] These policy functions are plotted from the perspective of a given firm, as functions of this firm's technological lead relative to its competitor(s), where a negative number means that the firm is lagging relative to its competitor.

In a two-superstar industry, the incentives to advertise are the highest when the firms are close to being neck-to-neck, and remain high when one firm has a slight lead. For larger leads, incentives decline. Indeed, when one of the firms is leading by a large gap, its incentives to advertise are relatively low because the firm does not gain too much additional demand relative to its competitor. The policy function in industries with three superstars exhibits a similar pattern, with advertising incentives increasing in the technological lead, and declining (though only slightly) when the firm is far ahead of both of its competitors.

Figure F.1 in the Appendix shows the corresponding policy functions for innovation, exhibiting a similar feature: firms innovate the most when they are close to being neck-to-neck, and innovation incentives decrease with higher technological gap with their competitors.

Advertising, innovation, and market share. As our main quantitative exercises will relate to the effects of advertising policy through endogenous responses in innovation, advertising, and market structure, we must also make sure that the model can reproduce the empirically observed relationships between these variables.

[21] Unlike the Kimball specification used in Edmond et al. (2023), our model allows for non-constant superelasticity across relative firm size.

[22] Since we assume that there exists a maximum technology gap $\bar{n}$ between any two superstars, some states on the right panel of Fig. 2 are impossible, which is why the policy function is displayed as a strip in the space of states.

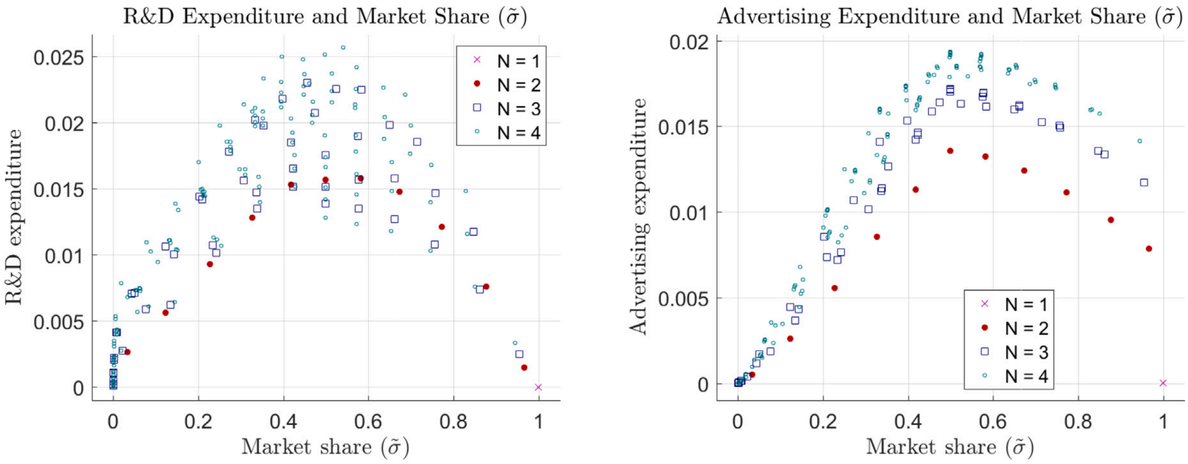


Fig. 3. R&D expenses, advertising, and firm market shares.
Notes: This figure displays firm-level R&D (left panel) and advertising (right panel) expenditures as functions of the firm's market share relative to other superstars in its industry (i.e., $\tilde{\sigma}$ defined in Eq. (13)). Each marker in these figures corresponds to the choice of a firm given an industry state, ranging from $N = 1$ to $N = 4$ superstars per industry.

Fig. 3 shows firm-level R&D (left panel) and advertising (right panel) expenditures in the model as functions of the firm's market share relative to other superstars in its industry (i.e., $\tilde{\sigma}$ defined in Eqs. (13)).[23] Each marker in these figures corresponds to the choice of a firm given an industry state, ranging from $N = 1$ to $N = 4$ superstars per industry. The figure shows that the model generates, within all industries, an inverted-U shaped relationship between a firm's innovation and advertising efforts and its share of sales in its industry (recall that the intercept and top point of both of these curves were targets of the calibration). Note that the inverted-U relationships continue to hold even within industries with the same number of superstars, which is also true in the data.[24]

5. Counterfactual experiments

We are now ready to tackle the central question of our paper: how does advertising affect social welfare through its static and dynamic interactions with innovation, competition, and market structure, as well as macroeconomic aggregates such as the level of output, markups, the labor share, and long-run economic growth? Further, what are the implications for policy intervention?

5.1. The macroeconomic effects of shutting down advertising

As a first pass, we conduct a counterfactual experiment in which we shut down advertising completely. In particular, we compare our baseline calibration with a stationary equilibrium in which superstar firms find it prohibitively costly to invest in advertising.[25]

5.1.1. The dynamic impact on macroeconomic aggregates

Table 2 reports how key macroeconomic aggregates change as a result of our experiment. We can first notice that the R&D share of GDP and economic growth increase when advertising is shut down. There are several forces at play regarding the relationship between aggregate advertising and R&D, as both R&D and advertising can be used by firms to shift demand away from competitors and towards their products. On the one hand, advertising allows firms to magnify the return on their innovation, hence increasing the incentives to perform R&D. From this point of view, advertising and R&D can be seen as complements. On the other hand, when firms cannot advertise, they lose one potential tool to differentiate their products from those of their competitors, and might invest more in the remaining tool – R&D – making advertising and R&D substitutes.

Our results in Table 2 suggest that the second effect dominates, and that R&D and advertising are substitutes at the aggregate level in general equilibrium: on average, innovation by superstars increases in response to shutting down advertising.[26] This result is

[23] Note that we plot advertising and R&D expenditures as a function of relative sales among superstars alone, and not all firms. The superstars in the model are mapped to publicly-traded US firms in the data (Compustat). The relative market share among superstars in the model is therefore mapped to the relative market share in the firm's SIC4 industry in Compustat.

[24] Recall that the set of controls in our regressions (1a)–(1b) includes the number of firms in the industry.

[25] In this case, the perceived quality of every single variety is equal to one, $\hat{\omega}_{ij} = 1$. This experiment is equivalent to the limiting case of our model in which the cost scale parameter of advertising χ_a goes to infinity.

[26] Under different parameter values, advertising and R&D can be complements rather than substitutes in our model. Therefore, this is a quantitative result rather than a theoretical implication of the framework.

Table 2
Advertising shutdown: The dynamic impact on macroeconomic aggregates.

	Benchmark	Advertising shutdown	% change
Growth rate	2.201%	2.273%	3.26%
R&D/GDP	2.467%	2.613%	5.92%
Advertising/GDP	0.022	0.000	−100.00%
Average markup	1.342	1.254	−6.58%
Std. dev. markup	0.442	0.340	−23.12%
Labor share	0.638	0.663	3.84%
Average profitability	0.136	0.126	−7.56%
Average leader relative quality	0.510	0.449	−11.84%
Std. dev. leader relative quality	0.164	0.144	−11.92%
Superstar innovation	0.339	0.394	16.17%
Small firm innovation	0.096	0.112	16.44%
Output share of superstars	0.431	0.422	−2.12%
Average superstars per industry	2.864	3.264	13.99%
Mass of small firms	1.000	1.328	32.84%
Initial output	1.159	1.105	−4.63%

Notes: This table presents the changes in the relevant macroeconomic aggregates under the advertising shutdown compared to the baseline economy.

in line with the empirical findings in Cavenaile and Roldan-Blanco (2021), providing an additional out-of-sample validation test for the model. Interestingly, small firms also raise their investment in R&D when advertising is shut down. This can be linked to results that we will further discuss in Section 5.1.2, where we argue that advertising shifts market shares from small to large superstars. As a result, the absence of advertising leads to a higher value of small superstar firms, and hence an increase in the incentives for small firms to perform R&D in order to become superstars themselves. Overall, shutting down advertising raises economic growth by 3.26% of its baseline value.

In addition, advertising also affects business dynamism. As advertising affects the value of small firms in the economy, it also changes the investment behavior of entrepreneurs. When advertising is shut down, entrepreneurs' investment rate increases and the mass of small firms in the economy goes up by 32.8%. In other words, advertising decreases business dynamism along two dimensions. First, it slows down the number of new small firms that are created and, second, it decreases the rate at which new superstars emerge. Shutting down advertising, on the other hand, levels the playing field, favoring smaller firms over the larger superstars.[27] Conversely, one could interpret this finding as advertising playing the role of a barrier to entry of new firms.

Firms in our model use advertising to shift demand towards their product and away from their competitors, and thereby charge higher markups. As a result, shutting down advertising also leads to a significant decrease in markups. The average net markup decreases from 34% to 25%. In other words, advertising is found to be responsible for roughly one quarter of the average net markup observed in the calibrated equilibrium, whereas the remaining three quarters are attributable to productivity heterogeneity and the love for variety of the consumers. The standard deviation of firm-level markups also falls by 23.1% of its value, implying that advertising is responsible for one quarter of the empirically observed dispersion in markups. The decrease in the average markup is accompanied with a decline in the profitability of superstar firms by 7.56%, and a rise in the labor share by 3.84% of its value.

While these changes seem beneficial for social welfare at first glance, the effects of shutting down advertising on static and dynamic allocative efficiency are found to be quite nuanced, which we investigate next.

5.1.2. The impact on static misallocation of resources

We have just shown that shutting down advertising leads to a significant decrease in the average markup and its dispersion. One might therefore be tempted to expect an increase in allocative efficiency. Interestingly, we find that shutting down advertising reduces static allocative efficiency, decreasing the level of output by 4.63% of its value.

This result owes to two effects, both working in the opposite direction compared to the change in markups. First, advertising is found to help reallocate production from less productive superstars (low q_{ij}) to more productive ones (high q_{ij}). Since more productive firms advertise more overall, demand shifts towards their products. Due to the oligopolistic competition under a CES aggregator as in Atkeson and Burstein (2008), more productive firms charge higher markups in equilibrium and produce too little relative to the efficient allocation. This means that, statically, the economy with advertising allocates inputs more efficiently, because advertising partially alleviates the misallocation of resources inherent to this environment with oligopolistic competition.

Second, advertising directly affects products' perceived quality ($\hat{\omega}$), and hence, welfare. In equilibrium, advertising is such that the perceived quality of large superstars (whose products are consumed more) is magnified compared to smaller superstars, which further amplifies the gains from production. Combined together, the reallocation of resources towards more efficient firms, and the relative amplification of the perceived quality of these firms, work against the effects of higher markups, implying that advertising helps improve static allocative efficiency on the net.

[27] This effect is heterogeneous across firms, and there are winners and losers. See Appendix B.4 for a discussion.

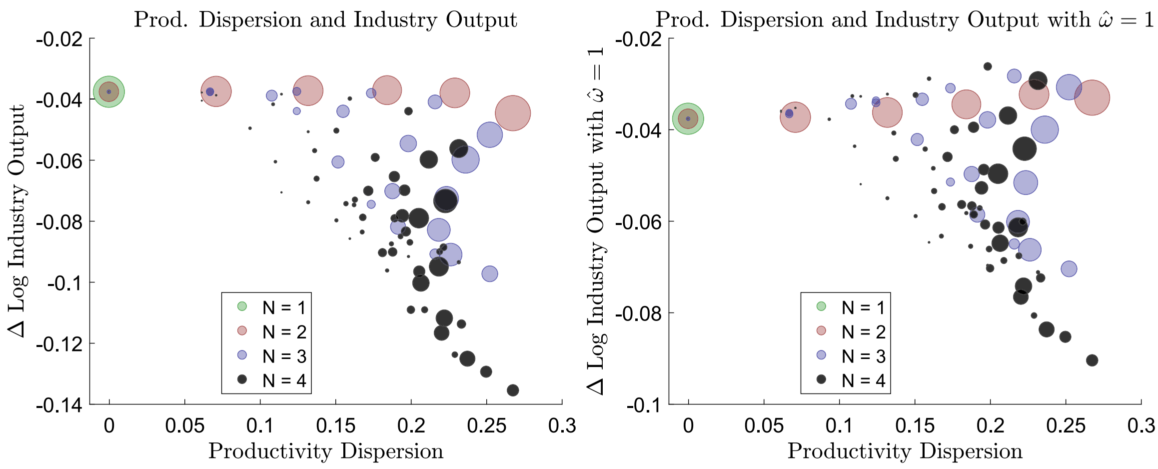


Fig. 4. Change in industry output by productivity dispersion.
Notes: This figure depicts how shutting down advertising affects allocative efficiency across different industry states. The left panel displays the change in industry output as a function of productivity dispersion in the industry. The right panel shows the change in industry output evaluated at a fixed perceived quality ($\hat{\omega}_{ij} = 1$) such that the change in industry output is solely due to changes in quantities produced. Each circle represents an industry state, with the color of the circle denoting the number of superstars in the industry, and the size of the circle indicating the share of that industry state in the baseline invariant distribution.

Fig. 4 shows that most of the allocative efficiency gains are due to the first effect. The left panel displays the change in log industry output as a function of productivity dispersion in the industry — the combined result of changes in the labor allocation between firms and the change in their perceived quality. In contrast, the right panel shows how much of the change in log industry output is due to labor reallocation alone. In particular, it shows the change in log industry output evaluated at a fixed perceived quality, $\hat{\omega}_{ij} = 1$, so that these changes are only due to quantities produced. Each circle represents an industry state, with the color of the circle denoting the number of superstars in the industry, and the size of the circle indicating the share of that industry state in the baseline stationary distribution. Comparing both panels of Fig. 4, changes in log industry output are very similar, which shows that labor reallocation between firms explains a large share of the reduction in misallocation.

We can also notice from Fig. 4 that the reallocation of production between firms, improving static allocative efficiency, is stronger in industries where the dispersion in terms of productivity is larger. Intuitively, it is in industries with more productivity dispersion that the potential gains to production reallocation between firms are the largest.

5.1.3. Short-run versus long-run effects on welfare

How do these various macroeconomic effects translate into social welfare? To address this question, we decompose the results described above into their static and dynamic parts. The static part owes to changes in observables for a given distribution over industry states. The dynamic effect is due to the endogenous response of firms in terms of R&D investment when advertising is shut down, which leads to a change in the distribution over industry states.[28]

Specifically, the model allows for an analytical decomposition of the change in welfare, $\Delta W \equiv W^B - W^A$, between any two stationary equilibria A and B as follows (see the details in Appendix C.4):

$$\Delta W = \frac{1}{\rho}\Bigg[\underbrace{-\Delta \ln w^{rel} + \Delta\left(\sum_{\Theta} f(\Theta)\mu(\Theta)\right)}_{\Delta \ln(Y_0)} + \Delta \ln\left(\frac{C_0}{Y_0}\right)\Bigg] + \frac{1}{\rho^2}\Delta g \tag{24}$$

The first term inside the square brackets reflects the change in the relative wage, and the second term relates to changes in the relative industry output of superstar firms. These two terms collectively represent the change in welfare due to the change in the initial level of output. The third term captures changes in the consumption share of GDP. The last term in the equation captures how the differential in the growth rates between the two economies is translated to changes in welfare.

Table 3 shows how each of these margins is affected by shutting down advertising. Overall, we obtain a welfare loss: the consumption-equivalent welfare change (CEWC) is –0.86%.[29] The first two columns in the table report the static effect of advertising

[28] In Appendix E, we state the social planner problem and derive closed-form solutions for the static part. This helps us identify misallocation of labor from markups as the source of static inefficiencies in the decentralized equilibrium.

[29] Consumption-equivalent welfare is defined as the compensation in lifetime consumption that the representative household from one economy requires to remain indifferent between consuming in this economy versus consuming in the counterfactual economy. This welfare measure is provided in equation (C.20) of Appendix C.4.

Table 3
Advertising shutdown: Short-run vs. long-run effects on efficiency.

	Static		Static+New Dist.		Dynamic	
	ΔW	CEWC	ΔW	CEWC	ΔW	CEWC
Relative wage	−0.883	−3.47%	−0.942	−3.70%	−0.942	−3.70%
Output of superstar firms	−0.618	−2.44%	−0.242	−0.96%	−0.242	−0.96%
Consumption/output	0.573	2.32%	0.573	2.32%	0.520	2.10%
Output growth	0.000	0.00%	0.000	0.00%	0.448	1.81%
Total	−0.927	−3.64%	−0.612	−2.42%	−0.217	−0.86%

Notes: This table shows the decomposition of the changes in social welfare between our baseline calibration and our counterfactual economy without advertising.

on welfare, i.e., fixing the distribution over industry states, $\mu(\Theta)$, and the level of R&D, at their baseline levels. Statically, shutting down advertising results in a large welfare loss of 3.64%, which comes from the resulting increase in the relative wage and decrease in the output of superstar firms. As discussed in Section 5.1.2, shutting down advertising reduces static allocative efficiency, which results in a welfare loss. The third and fourth columns in Table 3 display what happens to welfare if we further let the distribution adjust but still keep R&D and growth fixed. In that case, the welfare loss from shutting down advertising is smaller, at 2.42%. This is due to the fact that the industry state distribution shifts towards industries in which superstars produce more. As a result, the total output of superstars increases which results in welfare gains. On the other hand, the relative wage further increases.

Finally, the last two columns of Table 3 show the full results including the dynamic effects due to changes in R&D investment.[30] Shutting down advertising raises the consumption-to-output ratio as a result of changes in total R&D and advertising expenses. In addition, the growth rate of the economy increases. Overall, these dynamic effects further offset some of the static welfare losses. All in all, static losses are larger than dynamic gains, resulting in a total welfare loss of 0.86% in consumption equivalent terms when advertising is shut down.

5.2. Should we tax or subsidize advertising?

Our analysis in the previous section finds that advertising, despite its various negative effects, helps rather than hurts efficiency, albeit by a smaller margin than what we would find if the dynamic effects were ignored. This raises questions for policy intervention: should advertising be taxed or subsidized?

Table F.4 in the Appendix reports the results of comparing stationary equilibria across different values of advertising taxes and subsidies.[31] In line with the results of our shutdown experiment, higher taxes (subsidies) on advertising are associated with a reduction (increase) in advertising expenditures, and an increase (decrease) in innovation and aggregate productivity growth. Taxing advertising also results in a decrease in the average markup and its dispersion and in an increase in the labor share. At the same time, raising taxes also decreases the level of initial output as static allocative efficiency worsens. The decrease in advertising expenditures along with the lump-sum rebate of the tax results in an increase in initial consumption at low levels of the tax rate. As the tax rate keeps increasing, the decrease in initial output due to losses in static allocative efficiency dominates, and initial consumption starts decreasing.

Interestingly, even though shutting down advertising decreases welfare, the tax rate that maximizes welfare is positive and quite high, at 62.9% (see Figure F.2).[32] How does one reconcile these seemingly contradictory findings? The answer lies in understanding how taxation differs from a complete shutdown. Higher taxes on advertising expenses discourage firms from investing resources into advertising, resulting in both direct gains in the aggregate consumption share, and indirect gains from improved incentives for innovation and growth. However, taxes do not cause as large a drop in static allocative efficiency as a complete shutdown would: while the overall spending on advertising declines, more productive superstars still continue to spend more on advertising. Therefore, the positive effects of advertising due to the more efficient input allocation are still present even under high tax rates. In other words, the taxes reduce the excessive spending on advertising due to the “rat race” between the superstars, while still largely preserving the within-industry distribution of market shares in equilibrium. This makes advertising an ideal candidate for taxation.

In most advanced economies including the United States, advertising expenses are not taxed, as discussed in the introduction. Our quantitative analysis suggests that advertising is a useful activity insofar as it improves static allocative efficiency through a reduction

30 Though these welfare numbers are calculated by comparing the two stationary equilibria, we could also feasibly solve for the non-stationary equilibria which includes the transition from the old stationary equilibrium to the new one. This would result in welfare losses between what we calculate for the static and dynamic effects in Columns 2 and 6 of Table 3. This is because, during the transition, the industry state distribution $\mu_t(\Theta)$ and the mass of small firms m_t take time to converge to their new steady state values, which delays the positive effect of higher aggregate growth on welfare, whereas the static impact of advertising is instantaneous.

31 We focus on linear taxes and subsidies. The revenues from taxes are rebated back to the consumers, and subsidies are financed through lump-sum taxes. Note that reported tax and subsidy rates correspond to the share of total advertising-related expenses that are collected as tax or paid as subsidies by the government.

32 This tax is associated with a 0.64% increase in growth, a 2.22% increase in superstar innovation, a 6.43% reduction in the average net markup and 5.51% reduction in markup dispersion, a 0.95% increase in the labor share, a 1.44% reduction in initial output, a 4.15% increase in the mass of small firms, and an overall increase in welfare of 0.52%. Subsidies, on the other hand, only serve to reduce welfare.

in the misallocation of resources. However, the same useful effects can largely be attained under relatively high linear taxes, while eliminating most of the excessive spending that arises due to its “rat race” nature. Given that most taxes that governments levy to finance government spending unambiguously reduce efficiency rather than boost it, taxing advertising seems like a great alternative, which can be used to raise a significant amount of revenue – 1.12% of GDP under the optimal tax rate – while simultaneously improving dynamic efficiency.[33]

While the optimal level calculated at 62.9% may seem rather high, this is well within the range of taxes that European countries levy on petroleum products, which create a large dead-weight loss as well as increase transportation costs. We should also highlight that moderate advertising tax rates can still reap most of the benefits the optimal tax rate delivers. For instance, as the second column of Table F.4 shows, a modest 25% tax rate can still deliver 58.3% of the consumption-equivalent welfare gains that the optimal tax rate of 62.9% provides. In such a world of second-bests, taxation of advertising expenditures seems to be an idea well worth investigating, all the more so given that advertising expenditures are found to be very inelastic to the taxes levied.[34]

6. Robustness checks

In this section, we examine whether our main results – the positive effect of advertising on social welfare and the optimal linear taxation of advertising – are robust to some extensions and modifications of our baseline setup. We briefly discuss these robustness checks next and relegate their details to Appendix D.

6.1. *Ex-ante versus ex-post preferences and deceptive advertising*

One potential concern highlighted in the literature when evaluating the welfare effect of persuasive (taste-shifting) advertising relates to whether welfare should be computed using ex-ante or ex-post preferences (see, e.g., Dixit and Norman (1978) or Benhabib and Bisin (2011)).

In our baseline experiments, we evaluate welfare assuming that advertising influences consumers’ welfare in the same way that it influences revealed preferences from consumer demand. To break this link, we propose an extension of our model in which we allow for advertising to be (partly) deceptive: every instant, advertising in every industry turns out to be (unexpectedly) purely manipulative with probability $\delta \in [0,1]$. In this event, welfare is evaluated using the equilibrium allocation, but imposing that $\hat{\omega}_{ijt} = 1$ for all products i in that industry.[35] The case with $\delta = 0$ corresponds to our baseline model where ex-ante and ex-post preferences coincide, whereas $\delta = 1$ implies that advertising is fully deceptive. Consequently, δ parametrizes how severe the deceptive advertising problem is in the overall economy. Under this extended economy, none of the positive implications regarding the competitive equilibrium change, since purchases are still made according to the demand shifters $\hat{\omega}_{ijt}$ that prevail in the baseline model. Only the (normative) welfare calculation is altered.

Our results indicate that as advertising gets closer to being purely deceptive (i.e., δ increases), welfare losses from shutting down advertising decrease and can eventually turn into welfare gains. However, the optimal tax rate is still below 100%, and there is still a role for advertising to fulfill, thanks to its property of alleviating static misallocation.[36] That is, advertising is still socially useful even if it is purely deceptive.

6.2. *Non-combative advertising*

In the baseline model, as per Eq. (7), an increase in a firm’s advertising efforts decreases the perceived quality of every other product in the industry, other things equal. To relax this assumption, we can more generally assume that the perceived quality of variety i is given by

$$\hat{\omega}_{ijt} \equiv \frac{1+\omega_{ijt}}{\Lambda + \frac{1-\Lambda}{N_{jt}}\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \tag{25}$$

where $\Lambda \in [0,1]$ is a parameter that governs the degree of advertising combativeness across firms. When $\Lambda = 0$, we return to our baseline model. When $\Lambda = 1$, we have $\hat{\omega}_{ijt} = 1+\omega_{ijt}$, so that a firm’s advertising does not directly affect the perceived quality of other products.

Appendix D.2 derives the equilibrium conditions of this alternative model. To prove the robustness of our results, we perform two robustness checks. In the first check, we pick the extreme value of $\Lambda = 1$, re-calibrate the model under this choice, and repeat the experiments (see Tables F.6 and F.7). The extended model with $\Lambda = 1$ makes advertising more useful from a social perspective, and therefore the welfare cost of shutting down advertising is now much higher at 5.38%. This large increase primarily owes to the 80% higher impact of the shutdown on initial output, due to the increased direct benefit of advertising on welfare. In both models,

[33] In 2019, the tax-to-GDP ratio of the United States was 25.5%. This means optimal advertising taxes could raise 4.39% of the tax revenue already being collected through distortionary taxes.

[34] Sensitivity analysis reveals that the main parameter that determines the elasticity of total advertising to its tax rate is the advertising cost convexity parameter ϕ_a, the value of which is pinned down by targeting the top point of the inverted-U relationship between advertising and market share.

[35] In such instances, products of firms with $\hat{\omega} > 1$ are revealed to be “overhyped”, and those with $\hat{\omega} < 1$ are revealed to be “hidden gems”. In retrospect, the consumers would have preferred to purchase less of the former and more of the latter, but their purchases are already made at the time welfare is evaluated.

[36] See Appendix D.1 for more details.

shutting down advertising affects all economic quantities of interest in the same direction, although exact magnitudes vary. We also find that it is still optimal to tax advertising rather than subsidize it. The optimal tax rate is now 28.6%.

For the second robustness check, we re-calibrate the model and include the value of Λ as an internally calibrated parameter. To pin down its value, we target the ratio of the regression coefficients on own advertising to the regression coefficient on competitors' advertising, i.e., β_3/β_4 from regression (2).[37] We find a calibrated value of $\Lambda = 0.0199$ (see Table F.8), which is very close to zero as assumed in our baseline model. As a consequence, the results of our counterfactual experiments remain mostly unaffected (Table F.9).

6.3. Other robustness checks

Two more robustness exercises complete our sensitivity analysis. First, we solve our model assuming that firms compete in prices (à la Bertrand) and not quantities, calibrate this alternative model, and repeat our quantitative experiments. Appendix D.3 provides the details of how the best responses of firms and the static equilibrium conditions in each industry change. The results can be found in Tables F.10 and F.11. The advertising shutdown experiment reveals once again that advertising and innovation are substitutes at the aggregate level. Shutting down advertising still boosts innovation, business dynamism, economic growth, and the labor share, and average markups and markup dispersion go down. Similarly, the shutdown adversely affects initial output, as it increases static misallocation across superstars. However, in terms of welfare, the dynamic gains slightly dominate the static losses this time. Despite a minor gain in consumption-equivalent welfare after an advertising shutdown, the optimal advertising tax experiment reveals that taxing advertising heavily is still preferable to shutting it down altogether.

Second, we perform robustness checks on the maximum number of productivity steps between any two firms within the industry ($\bar{n}$) and on the maximum number of superstar firms per industry ($\bar{N}$). These parameters were set to $\bar{n} = 5$ and $\bar{N} = 4$ in our baseline calibration. We now set $\bar{n} = 6$ and $\bar{N} = 5$, respectively, and recalibrate the model. Calibration results can be found in Tables F.12 and F.14, and results of the counterfactual shutdown experiments are reported in Tables F.13 and F.15. The results of these experiments are very similar to those obtained in our baseline calibration.[38]

7. Conclusion

Firms routinely make intensive use of innovation and advertising in order to alter their process efficiency and the perceived quality of their products, allowing them to shift consumer demand toward themselves and gain market share in their industry. At the aggregate level, these two forms of intangible investments account for a large share of GDP in the United States. Yet, the interaction between them and their implications for economic growth and social welfare remain understudied in the literature.

In this paper, we have uncovered a non-monotonic empirical relationship between innovation, advertising and firms' relative sales, hinting at the idea that these different forms of intangibles might be used strategically by firms to gain market share against their competitors. Motivated by this evidence, we have built a unified framework to study the interaction between advertising and innovation in a heterogeneous-firm model in which market structure (i.e., the number of large firms and their relative productivities), markups, and growth are all endogenous.

When calibrated to match our micro-level facts, our model predicts advertising to have important quantitative implications for macroeconomic aggregates. Since advertising and innovation are substitutes in the calibrated model, shutting down advertising improves the incentives of firms to innovate instead, which boosts economic growth. However, we find that advertising also helps improve static allocative efficiency through reallocating resources towards more efficient firms, and its shutdown therefore reduces static efficiency. On the net, static losses are larger than dynamic gains, resulting in a total welfare loss of 0.86% in consumption-equivalent terms, implying that advertising is a useful economic activity from a social perspective.

In spite of this result, we find that advertising should be taxed, and that the welfare-maximizing linear advertising tax would lead to a 0.64% increase in growth, a 6.43% reduction in the average net markup, and a 0.95% increase in the labor share, for an overall increase in welfare of 0.52%. In other words, despite its positive effects on static allocative efficiency, a linear tax levied on advertising can still improve welfare since it can reduce the excessive spending that is due to the "rat race" nature of advertising. The distortion-free tax revenue raised at 1.12% of GDP is an added bonus that can lead to further welfare gains through reduced reliance on other sources of taxation that are more distortionary for economic activity. These positive taxation results are found to be robust across all our model extensions. Our results therefore provide a justification for the recent efforts by policymakers to impose taxes on (digital) advertising.[39] We hope that our analysis and, more broadly, our new quantitative framework, prove useful in coming up with welfare-improving policy interventions regarding advertising and innovation.

Data availability

The authors do not have permission to share data.

[37] A higher value of Λ lowers the direct effect of a firm's advertising on the perceived quality of other firms, reducing the degree of "combativeness".

[38] See Appendix D.4 for a discussion of the results.

[39] There could also be other unmodeled mechanisms that could push the results towards even higher optimal advertising tax rates, such as agency frictions between the manager and the shareholders, short-termism, and over-advertising to diminish the impact of negative demand shocks (such as a scandal regarding the firm or its products). Extending our framework along these dimensions could be a good avenue for future research.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmoneco.2024.103683.

References

Afrouzi, H., Drenik, A., Kim, R., 2023. Concentration, market power, and misallocation: the role of endogenous customer acquisition. Working Paper.
Aghion, P., Bergeaud, A., Boppart, T., Klenow, P.J., Li, H., 2023. A theory of falling growth and rising rents. Rev. Econ. Stud. 90 (6), 2675–2702.
Argente, D., Fitzgerald, D., Moreira, S., Priolo, A., 2023. How do firms build market share? the role of demand frictions. Amer. Econ. Rev.: Insights (forthcoming).
Arkolakis, C., 2010. Market penetration costs and the new consumer margin in international trade. J. Polit. Econ. 118 (6), 1151–1199.
Atkeson, A., Burstein, A., 2008. Pricing-to-market, trade costs, and international relative prices. Amer. Econ. Rev. 98 (5), 1998–2031.
Bagwell, K., 2007. The economic analysis of advertising. Handb. Ind. Organ. 3, 1701–1844.
Baslandze, S., Greenwood, J., Marto, R., Moreira, S., 2023. The expansion of varieties in the new age of advertising. Rev. Econ. Dyn. 50, 171–210.
Becker, G.S., Murphy, K.M., 1993. A simple theory of advertising as a good or bad. Q. J. Econ. 108 (4), 941–964.
Benhabib, J., Bisin, A., 2002. Advertising, mass consumption and capitalism. Working Paper, New York University.
Benhabib, J., Bisin, A., 2011. Social construction of preferences: Advertising. In: Benhabib, J., Bisin, A., Jackson, M. (Eds.), Handbook of Social Economics. pp. 201–220.
Butters, G.R., 1977. Equilibrium distributions of sales and advertising prices. Rev. Econ. Stud. 44 (3), 465–491.
Cavenaile, L., Celik, M.A., Perla, J., Roldan-Blanco, P., 2023. A theory of dynamic product awareness and targeted advertising. Working Paper.
Cavenaile, L., Celik, M.A., Tian, X., 2019. Are markups too high? Competition, strategic innovation, and industry dynamics. Working Paper.
Cavenaile, L., Roldan-Blanco, P., 2021. Advertising, innovation, and economic growth. Am. Econ. J.: Macroecon. 13 (3), 251–303.
Covarrubias, M., Gutiérrez, G., Philippon, T., 2020. From good to bad concentration? US industries over the past 30 years. NBER Macroecon. Annu. 34 (1), 1–46.
Crouzet, N., Eberly, J.C., 2019. Understanding weak capital investment: The role of market concentration and intangibles. Proceedings of the 2018 Jackson Hole symposium 87–148.
Danaher, P.J., Bonfrer, A., Dhar, S., 2008. The effect of competitive advertising interference on sales for packaged goods. J. Mar. Res. 45 (2), 211–225.
De Loecker, J., Eeckhout, J., Unger, G., 2020. The rise of market power and the macroeconomic implications. Q. J. Econ. 135 (2), 561–644.
De Ridder, M., 2024. Market power and innovation in the intangible economy. Amer. Econ. Rev. 114 (1), 199–251.
Dinlersoz, E., Goldschlag, N., Yorukoglu, M., Zolas, N., 2023. On the role of trademarks: From micro evidence to macro outcomes. Working Paper.
Dinlersoz, E.M., Yorukoglu, M., 2012. Information and industry dynamics. Amer. Econ. Rev. 102 (2), 884–913.
Dixit, A., Norman, V., 1978. Advertising and welfare. Bell J. Econ. 9 (1), 1–17.
Doganoglu, T., Klapper, D., 2006. Goodwill and dynamic advertising strategies. Quant. Mark. Econ. 4 (1), 5–29.
Dorfman, R., Steiner, P.O., 1954. Optimal advertising and optimal quality. Amer. Econ. Rev. 44 (5), 826–836.
Drozd, L.A., Nosal, J.B., 2012. Understanding international prices: Customers as capital. Amer. Econ. Rev. 102 (1), 364–395.
Dubé, J.-P., Hitsch, G.J., Manchanda, P., 2005. An empirical model of advertising dynamics. Quant. Mark. Econ. 3 (2), 107–144.
Edmond, C., Midrigan, V., Xu, D.Y., 2023. How costly are markups? J. Polit. Econ. 131 (7), 1619–1675.
Einav, L., Klenow, P.J., Levin, J.D., Murciano-Goroff, R., 2022. Customers and retail growth. Working Paper.
Feijoo Moreira, S., 2021. Provider-driven Complementarity and Firm Dynamics. Working Paper.
Fitzgerald, D., Haller, S., Yedid-Levi, Y., 2023. How exporters grow. Rev. Econ. Stud. 91 (4), 2276–2306.
Gouin-Bonenfant, É., 2022. Productivity dispersion, between-firm competition, and the labor share. Econometrica 90 (6), 2755–2793.
Gourio, F., Rudanko, L., 2014. Customer capital. Rev. Econ. Stud. 81 (3), 1102–1136.
Greenwood, J., Ma, Y., Yorukoglu, M., 2024. 'You will:' a macroeconomic analysis of digital advertising. Rev. Econ. Stud. http://dx.doi.org/10.1093/restud/rdae067, (forthcoming).
Gutiérrez, G., Jones, C., Philippon, T., 2021. Entry costs and aggregate dynamics. J. Monetary Econ. 124, S77–S91.
Ignaszak, M., Sedláček, P., 2022. Customer acquisition, business dynamism and aggregate growth. Working Paper.
Klein, M.A., Şener, F., 2023. Product innovation, diffusion and endogenous growth. Rev. Econ. Dyn. 48, 178–201.
Leone, R.P., 1995. Generalizing what is known about temporal aggregation and advertising carryover. Mark. Sci. 14 (3), 141–150.
Lind, J.T., Mehlum, H., 2010. With or without u? The appropriate test for a U-shaped relationship: Practitioners' corner. Oxf. Bull. Econ. Statist. 72 (1), 109–118.
Molinari, B., Turino, F., 2017. Advertising and aggregate consumption: A Bayesian DSGE assessment. Econ. J. 128 (613), 2106–2130.
Morlacco, M., Zeke, D., 2021. Monetary policy, customer capital, and market power. J. Monetary Econ. 121, 116–134.
Pearce, J., Wu, L., 2022. Brand reallocation, concentration, and growth. SSRN Working Paper.
Rachel, L., 2022. Leisure-enhancing technological change. Working Paper.
Roldan-Blanco, P., Gilbukh, S., 2021. Firm dynamics and pricing under customer capital accumulation. J. Monetary Econ. 118, 99–119.
Shapiro, B.T., Hitsch, G.J., Tuchman, A.E., 2021. TV advertising effectiveness and profitability: Generalizable results from 288 brands. Econometrica 89 (4), 1855–1879.
Terui, N., Ban, M., Allenby, G.M., 2011. The effect of media advertising on brand consideration and choice. Mark. Sci. 30 (1), 74–91.
Weiss, J., 2020. Intangible investment and market concentration. Working Paper.

Online Appendices:

Style Over Substance? Advertising, Innovation, and Endogenous Market Structure

Laurent Cavenaile[*] Murat Alp Celik[†] Pau Roldan-Blanco[‡] Xu Tian[§]

A Empirical Appendix

In this section, we provide details on data sets and variable construction, and report the results on the empirical firm-level relationships between innovation, advertising, and market share for public firms in the United States.

A.1 Variable Construction

Data Sources: We use the patent grant data obtained from NBER Patent Database Project which covers the years 1976-2006, and rely on Compustat North American Fundamentals for financial statement information of US-listed firms for the same years. Following a dynamic assignment procedure, we link the two data sets.[5]

Patent Citations: Our preferred measure of innovation is the number of citations a patent received as of 2006. We use the truncation correction weights devised by Hall *et al.* (2001) to correct for systematic citation differences across different technology classes and for the fact that earlier patents have more years during which they can receive citations (truncation bias). Our results are robust to using other innovation metrics, such as tail innovations, originality, and generality.[6] We calculate a firm-year level variable by calculating the average patent citations of all the patents the firm has applied for in a given year.

[*]University of Toronto Scarborough and Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON, M5S 3E6, Canada. Email: laurent.cavenaile@utoronto.ca

[†]University of Toronto, 150 St. George Street, Toronto, ON, M5S 3G7, Canada. Email: murat.celik@utoronto.ca

[‡]Banco de España and CEMFI, Calle Alcalá 48, 28014, Madrid, Spain. Email: pau.roldan@bde.es

[§]Terry College of Business, University of Georgia, 620 South Lumpkin Street, Athens, GA, 30602, USA. Email: xu.tian@uga.edu

[5]We have selected this time period based on data availability constraints. Our primary data source for innovation variables is the USPTO NBER Utility Patent Database, which accurately considers mergers and acquisitions, and the attribution of patents from subsidiaries to parent companies. This dataset only extends up to the year 2006. In line with the guidance provided by the authors (Hall *et al.* (2001)), we have opted not to include data from the last two years. This decision allows patents sufficient time to accumulate citations.

[6]See Section A.3 of Cavenaile *et al.* (2021) for the robustness checks using these alternative innovation metrics.

R&D Spending: Patent-based measures capture a successful innovation outcome. However, it might also be worthwhile to look at the amount of resources spent by a firm to conduct innovation regardless of success, as this captures the firm's intent. For this purpose, we use the R&D spending reported in Compustat (`xrd`). We replace missing values with zeroes, but the results are robust to dropping missing observations instead.

Advertising: We use the advertising spending reported in Compustat (`xad`) to measure advertising. It should be noted that this measure excludes in-house spending on marketing. Our focus is only on advertising purchased by firms from other firms, which this variable correctly captures.[7]

Relative Sales: We measure the relative sales of a firm among superstar firms in the same industry by dividing its sales by the total industry sales by all US public firms in that year in the same 4-digit SIC industry. The model counterpart of this variable is a superstar firm's market share among all superstars in the same industry (i.e., excluding the competitive fringe).

Control Variables: Our control variables include profitability, computed as $\texttt{oibdp}/\texttt{at}$; leverage, computed as $(\texttt{dltt}+\texttt{dlc})/\texttt{at}$; market-to-book ratio, computed as $(\texttt{csho}*\texttt{prcc\_c}+\texttt{at}-\texttt{ceq})/\texttt{at}$; log R&D stock, computed using an annual depreciation rate of 0.15; firm age; the coefficient of variation of the firm's stock price (from CRSP); the number of firms in the same 4-digit SIC industry; and a full set of year and 4-digit SIC industry fixed effects.

A.2 Regression Results

TABLE A.1: FIRM INNOVATION, ADVERTISING, AND RELATIVE SALES

	average patent citations	log advertising expenses
relative sales	0.629	6.260
	(0.095)***	(0.195)***
relative sales sq.	-0.623	-5.868
	(0.114)***	(0.255)***
R^2	0.15	0.73
N	104,911	37,779

Notes: Robust asymptotic standard errors reported in parentheses are clustered at the firm level. The sample period is from 1976 to 2004 at the annual frequency. All regressions control for profitability, leverage, market-to-book ratio, log R&D stock, firm age, the coefficient of variation of the firm's stock price, the number of firms in the industry, and a full set of year and SIC4 industry fixed effects. $^{***}=p<0.01,^{**}=p<0.05,^{*}=p<0.1$.

[7]Our empirical results are also robust to using selling, general, and administrative expenses instead (xsga), which includes in-house marketing spending, among other expenses.

Table A.2: Relative Sales, Own Advertising, and Others' Advertising

	relative sales	relative sales
own log advertising	0.072	0.071
	(0.003)***	(0.003)***
others' total log advertising	-0.072	
	(0.005)***	
others' avg. log advertising		-0.059
		(0.004)***
R^2	0.59	0.58
N	36,778	36,750

Notes: Robust asymptotic standard errors reported in parentheses are clustered at the firm level. The sample period is from 1976 to 2004 at the annual frequency. All regressions control for profitability, leverage, market-to-book ratio, log R&D stock, firm age, the coefficient of variation of the firm's stock price, the number of firms in the industry, and a full set of year and SIC4 industry fixed effects. $^{***} = p < 0.01, ^{**} = p < 0.05, ^{*} = p < 0.1$.

Table A.3: Firm Innovation, Advertising, and Relative Sales (Inverted-U Hypothesis Test)

	average patent citations	log advertising expenses
lower bound		
t-value	6.617	31.781
P>|t|	0.000	0.000
upper bound		
t-value	-4.237	-16.046
P>|t|	0.000	0.000

Notes: To further check the robustness of the inverted-U relationship between firm innovation, advertising, and relative sales, we test whether or not the slope of the fitted curve is positive at the start and negative at the end of the interval of the relative sales following Lind and Mehlum (2010). This table reports the hypothesis test results.

B Calibration Details and Additional Quantitative Results

B.1 Calibration Procedure

The model has 12 parameters to be determined: the innovation step size λ, the cost scale parameters for superstars and small firms (χ, ν), the corresponding cost curvature parameters (ϕ, ϵ), the relative productivity between the leader and the fringe ζ, the small firm exit rate τ, the entry cost scale ψ, the cost scale and curvature parameters in the advertising cost function (χ_a, ϕ_a), the elasticity among superstars' outputs within an industry η, and the elasticity between the superstars' output and the fringe's output γ. These 12 parameters are jointly calibrated via a moment-matching procedure, to match 14 moments in the data. The estimator is defined as the solution to the minimization of the weighted average distance between data and model moments.

B.2 Data Moments and Sources

We target the moments listed in Panel B of Table 1. In this section, we describe how we construct these data moments and provide the relevant data sources for each of these moments. All moments are calculated for the time period 1976-2004 using US data described in Appendix A.1.

1. **Growth rate:** To discipline output growth in our model, we obtain the annual growth rate of real GDP per capita from the US Bureau of Economic Analysis, and calculate the geometric average in our time period.

2. **R&D share of GDP:** The data for aggregate R&D intensity is taken from the National Science Foundation, which reports total R&D expenditures divided by GDP.[8]

3. **Average and dispersion in markups:** To discipline markups, we target the sales-weighted average markup and the sales-weighted standard deviation of markups found in De Loecker *et al.* (2020).

4. **Labor share:** We obtain the labor share estimates from Karabarbounis and Neiman (2013); in particular the time series for the corporate labor share (OECD and UN). For the capital share, we rely on the data from Barkai (2020). For both time series, we calculate the averages across all years for our sample. In our model, there is no capital. Therefore, the model-generated labor share $w^{rel}L = wL/Y$ corresponds to

[8] We target the aggregate R&D intensity for the US rather than relying on firm-level R&D intensity measures because such measures are available only for a selected sample of US firms.

the share of labor income among labor income plus profits. For comparability, we multiply this number by $(1-\kappa)$ where κ is the (exogenous) capital share, following Akcigit and Ates (2023).

5. **Firm entry rate:** In our model, firm entry rate is defined as the entry rate of new small firms. We obtain the data counterpart – the entry rate of new businesses – from the Business Dynamics Statistics (BDS) database compiled by the US Census Bureau.

6. **Relationship between innovation and relative sales, and advertising and relative sales:** We target the relationships between firm innovation and relative sales from regression (1a), and between firm advertising and relative sales from regression (1b). Innovation in the model is measured as the Poisson arrival event of quality improvement.

7. **Average profitability:** In the model, average profitability is calculated as the static profit flow minus advertising and R&D expenses divided by sales. In the data, it is defined as operating income before depreciation divided by sales (`oibdp`/`sale` in Compustat).

8. **Average and dispersion in leader relative quality:** We target the average relative quality of the leader in an industry, and its standard deviation across all industries. In the model, quality is known. In the data, we proxy quality by calculating the stock of past patent citations. The relative quality of the leader is defined as the quality of the leader divided by the sum of the qualities of the top four firms in an industry (SIC4 in the data).

9. **Advertising share of GDP:** The aggregate advertising expenses over GDP ratio is calculated based on the Coen Structured Advertising Expenditure Dataset, extracted from the McCann Erikson advertising agency.[9]

B.3 Identification

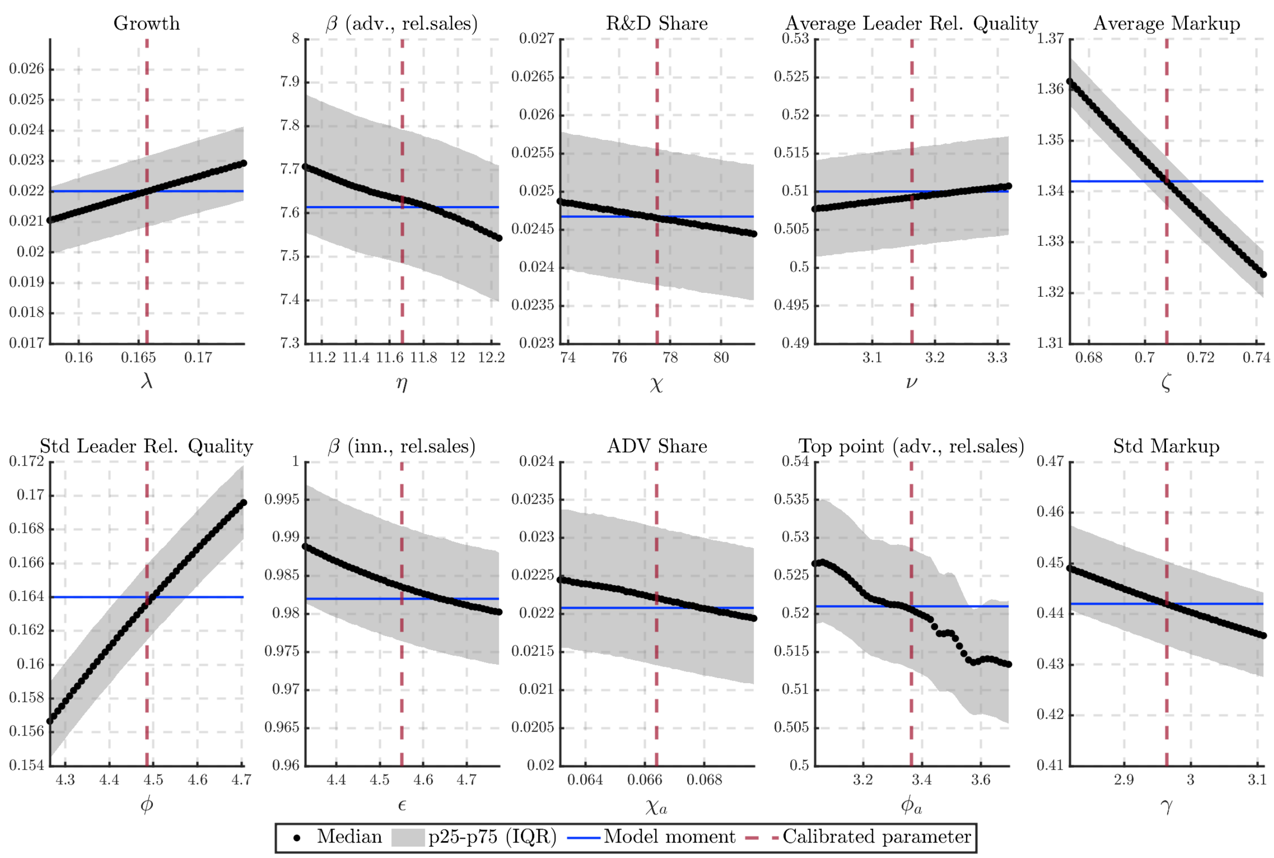
The model is highly nonlinear, and all parameters affect all the moments. Nevertheless, some parameters are more important for certain statistics. The success of the calibration requires that we choose moments that are sensitive to variations in the structural parameters. We now rationalize the moments that we choose to match.

To test for identification, we follow a method used by Daruich (2023). For each parameter-moment pair established in the text (see Table 1 for the summary), we allow

[9]The data is available at http://www.purplemotes.net/2008/09/14/us-advertising-expenditure-data/.

for quasi-random variation in all remaining parameters and solve the model for each such parameter configuration. In particular, first, we set wide bounds on each internally calibrated parameter. Then, we pick quasi-random realizations from the resulting hypercube using a Sobol sequence, which successively forms finer uniform partitions of the space. Finally, for each parameter combination, we solve the model and store the relevant moments. In total, we solve the model for about 1.8 million different parameter combinations. As a result, for each level of the identified parameter we obtain a whole distribution for the targeted moment.

FIGURE B.1: GLOBAL IDENTIFICATION TEST



Notes: This figure plots the results of our global identification test. For each moment-parameter combination, and across different values of the parameter of interest, we plot the median (black dotted line) and the inter-quartile range (gray shaded area) from the distribution that results from solving the model for 1.8 million quasi-random realizations of the remaining parameters. In each case, we also show the value of the parameter in the baseline calibration (red dashed vertical line) and the value of the moment predicted by the model (blue solid horizontal line).

In Figure B.1 we plot, for each parameter-moment pair, the median of this distribution (black dots) and the inter-quartile range (shaded area), together with the value of the parameter in the baseline calibration (red dashed vertical line) and the value of the moment predicted by the model (blue solid horizontal line). We consider that a parameter is well-identified by the moment when (i) the distribution changes across different values of the parameter; (ii) the rate of this change is high; and (iii) the inter-quartile range of the moment's distribution is narrow throughout the support for the parameter. Criterion (i) implies that the moment is globally sensitive to variation in the parameter; (ii) gives an idea of how strong this sensitivity is; and (iii) measures how much other parameters matter to explain variation in the moment. Because all the remaining parameters are not fixed but instead are varying in a quasi-random fashion within a wide support, this method gives us a global view of identification.

By these criteria, Figure B.1 shows that most of our parameters are overall well identified by the selected moments. In what follows, we offer a verbal intuition for this identification:

(i) The productivity step size parameter λ is mainly identified by the growth rate of aggregate output. A higher λ implies a higher increase in firm productivity upon successful innovation, which leads to higher output growth.

(ii) The linear term of the inverted-U relationship between advertising and relative sales and the standard deviation of markups are most helpful in identifying the elasticity of substitution between superstar firms η and the elasticity of substitution between superstar and small firms γ. Larger γ implies higher substitution between superstar and small firms, which leads to lower market power, profitability, and heterogeneity in markups across firms. Larger η implies higher substitution among superstars, which creates higher incentives for leading superstar firms to invest in advertising to shift demand and profits toward their products.

(iii) An increase in either superstar innovation cost scale parameter χ or small firm innovation cost scale parameter ν reduces the aggregate R&D share and average leader relative quality. Larger χ tends to reduce the innovation of superstar firms, narrowing the quality gaps between the industry leader and other superstar firms and leading to lower R&D spending on average. Higher ν increases the R&D cost of small firms, which reduces their innovation, leading to a reallocation of market share to superstar firms and a higher heterogeneity in qualities among superstar firms, increasing leader relative quality.

(iv) The relative productivity of small firms ζ is identified very precisely by matching the average markup. Lower ζ implies reduced competition from small firms and a within-industry market share reallocation to superstar firms, which generates a higher average markup and a lower labor share.

(v) As innovation policies in our calibrated model are below unity, an increase in the superstar R&D cost convexity parameter ϕ reduces the innovation cost, making it easier for large superstars to increase their technology gap, thereby increasing the dispersion in leader relative qualities. An increase in the small firm R&D cost convexity parameter ϵ, on the other hand, shifts the relative R&D spending of small firms vis-a-vis large ones, influencing the linear coefficient of the innovation-market share regression.

(vi) Intuitively, the three advertising-related data moments (the advertising share of GDP, and the linear term and top point of the hump-shaped relationship between advertising and relative sales) help us identify the two parameters governing the cost scale and curvature parameters in the advertising cost function, (χ_a, ϕ_a). While all the advertising related moments are affected by these two parameters, we find that the advertising share is most sensitive to changes in the advertising cost scale parameter χ_a, while the advertising cost curvature parameter ϕ_a affects the top point of the advertising to relative sales relationship most strongly.

Finally, two parameters, (τ, ψ), are identified directly:

(i) The exogenous small firm exit rate parameter τ is directly identified by targeting the entry rate of new businesses, since small firm entry rate equals small firm exit rate in a stationary equilibrium.

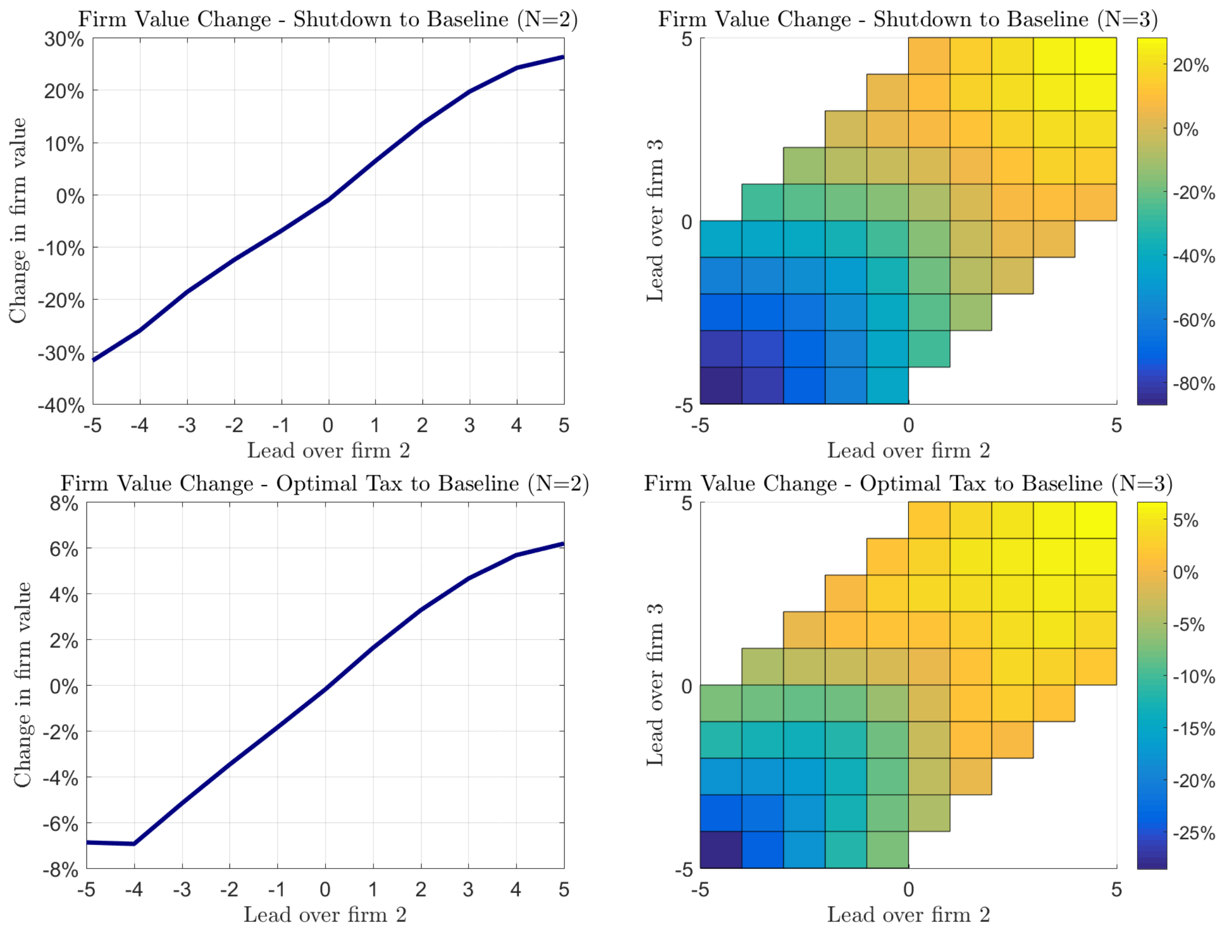
(ii) Given all other parameter values, the value of ψ is set to normalize the stationary measure of small firms m to one. Its exact value hinges on the average value of small firms, which itself is determined by the values of all other parameters. In particular, setting $m = 1$, we can rewrite equation (22) to get $\psi = \frac{1}{2\tau} \sum_{\Theta} v^e(\Theta)\mu(\Theta)$.

B.4 Additional Results: Heterogeneous Effects on Firms

Both the advertising shutdown and the optimal advertising tax experiments have heterogeneous effects on the firms in the economy, creating winners and losers. In this section, we investigate the heterogeneous firm-level effects of the two counterfactuals, and assess how much a firm's value and market share change between the two hypothetical economies and the calibrated baseline economy.

The top two panels of Figure B.2 show the change (in percentage terms) in firm value when moving from the BGP equilibrium with the advertising shutdown to the baseline economy. The left panel presents the results for 2-superstar industries, and the right panel does the same for 3-superstar industries, as a function of the technology gap between firms. The bottom two panels of Figure B.2 do the same for moving from the economy with the optimal advertising tax (62.9%) to the baseline economy. Finally, the four panels in Figure B.3 repeat the same exercise for market shares instead of firm value.

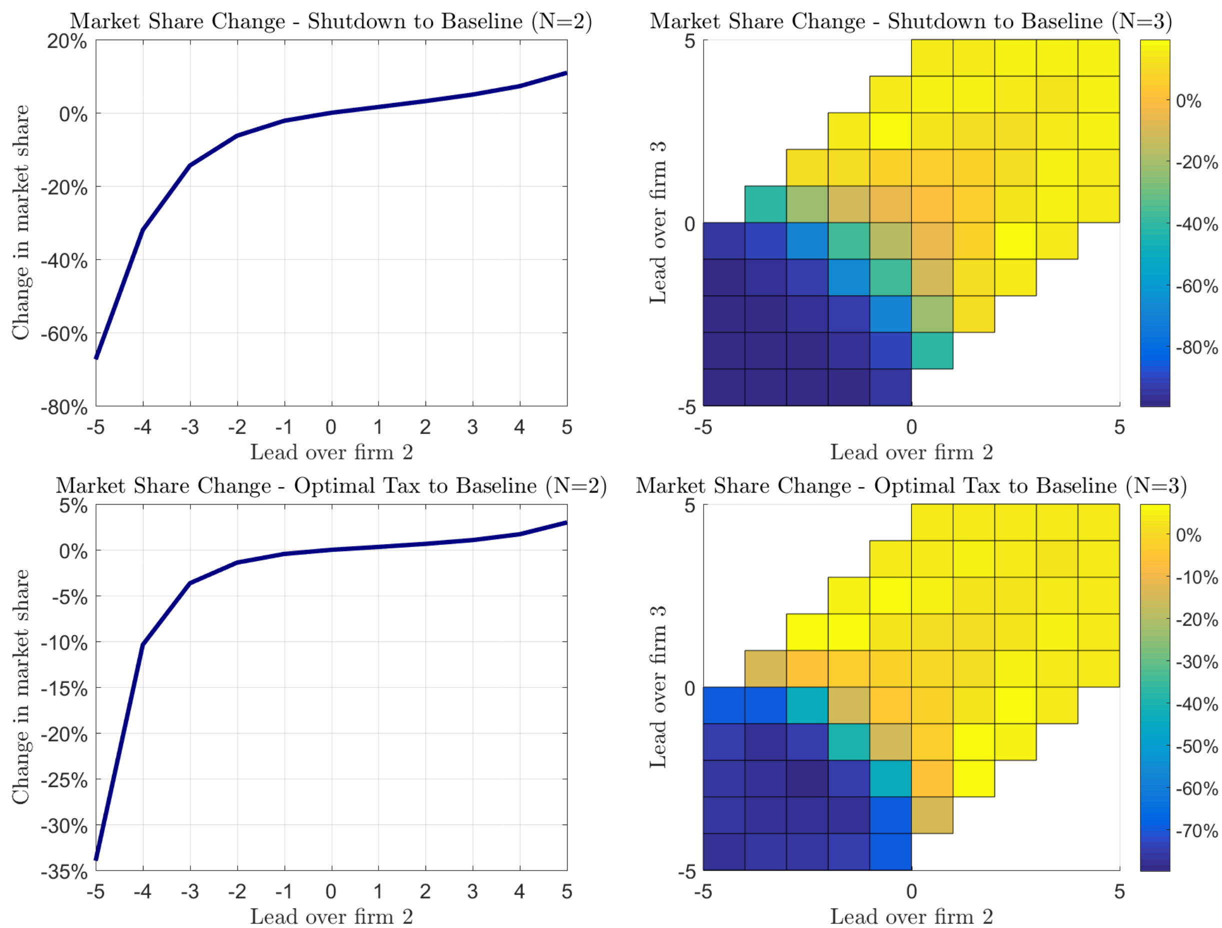
FIGURE B.2: CHANGE IN FIRM VALUE FROM SHUTDOWN AND OPTIMAL TAX



Notes: The top two panels show the change (in percentage terms) in firm value when moving from the BGP equilibrium with an advertising ban to the baseline economy without taxes for 2-superstar industries (upper-left panel) and 3-superstar industries (upper-right panel), as a function of the technology gap between firms. The bottom two panels show the change (in percentage terms) in firm value when moving from the BGP equilibrium with the optimal taxation level to the baseline economy without taxes for 2-superstar industries (bottom-left panel) and 3-superstar industries (bottom-right panel), as a function of the technology gap between firms.

As these figures show, both in terms of value as well as market shares, the less productive (laggard) superstar firms gain and the more productive (leading) superstar firms lose, both

FIGURE B.3: CHANGE IN MARKET SHARES FROM SHUTDOWN AND OPTIMAL TAX



Notes: The top two panels show the change (in percentage terms) in market share when moving from the BGP equilibrium with an advertising ban to the baseline economy without taxes for 2-superstar industries (upper-left panel) and 3-superstar industries (upper-right panel), as a function of the technology gap between firms. The bottom two panels show the change (in percentage terms) in market share when moving from the BGP equilibrium with the optimal taxation level to the baseline economy without taxes for 2-superstar industries (bottom-left panel) and 3-superstar industries (bottom-right panel), as a function of the technology gap between firms.

from shutting down advertising as well as from the introduction of taxes on advertising expenditures. Particularly, when moving from the economy with optimal taxes to the baseline economy, the market share of the most laggard firms declines by 34%, and their value goes down by around 7%. At the other end of the distribution, removing taxes would yield a 6% gain in firm value, and a 2.5% increase in market share, for the leading firms. Compared to the complete shutdown, the changes in value are roughly one-fifth as large for all firms when the optimal tax policy is implemented, relative to the baseline economy. The changes in market shares, by contrast, are quite asymmetric across firms within the same industry, with laggard firms losing a lot more (in percentage terms) than what leading firms gain when the taxes are removed. All in all, taxation redistributes both value and market share away from top superstar firms toward more laggard firms.

The increase in the value of laggard superstar firms also contributes to an increase in the value of small firms in the competitive fringe across the board, as evidenced by the 32.8% increase in the mass of small firms in the shutdown experiment and the 4.15% increase under optimal advertising taxes. This is because entrepreneurs react to the increase in small firm value by founding more new businesses, increasing business dynamism, small firm innovation, and consequently, leading to a higher number of superstar firms on average across industries. Entrepreneurial rents are therefore also magnified.

C Theory Appendix

C.1 Markov Perfect Equilibrium Definition

Given initial conditions, the unique Markov Perfect Equilibrium of this economy is defined by a set of allocations $\{C_t, Y_t, y_{ijt}, y_{ckjt}\}$, policies $\{l_{ijt}, l_{ckjt}, \omega_{ijt}, z_{ijt}, X_{kjt}, e_t\}$, prices $\{p_{ijt}, p_{cjt}, w_t, r_t\}$, the number of superstars in each industry N_{jt}, a measure m_t of small firms, and a set of vectors $\{\mathbf{n}_{ijt}\}$ which denote the relative productivity distance between firm i and every other firm in the same industry j at time t, such that, for all $t \geq 0$, $j \in [0, 1]$, $i \in \{1, ..., N_{jt}\}$:

(i) Given prices, final good producers maximize profit.

(ii) Given $\mathbf{n}_{ij}$ and N_{jt}, superstars choose y_{ijt} and ω_{ijt} to maximize profit.

(iii) Given prices, small firms in the competitive fringe choose y_{ckjt} to maximize profit.

(iv) Superstar firms choose innovation policy z_{ijt} to maximize firm value.

(v) Small firms choose innovation policy X_{kjt} to maximize firm value.

(vi) Entrepreneurs choose e_t to maximize entrepreneurial rents.

(vii) The real wage rate w_t clears the labor market.

(viii) Aggregate consumption C_t grows at rate $r_t - \rho$.

(ix) The aggregate resource constraint is satisfied:

$$Y_t = C_t + \int_0^1 \sum_{i=1}^{N_{jt}} \chi z_{ijt}^{\phi} Y_t \, dj + \int_0^1 \sum_{i=1}^{N_{jt}} \chi_a \omega_{ijt}^{\phi_a} Y_t \, dj + \int_0^1 m_t \nu X_{kjt}^{\epsilon} Y_t \, dj + \psi e_t^2 Y_t \tag{C.1}$$

The aggregate resource constraint states that the final output is used for consumption, superstars' R&D, superstars' advertising costs, R&D costs for small firms and entry costs.

C.2 Derivation of the Static Equilibrium Conditions

Inverse Demand Functions The final good is produced competitively. The cost minimization problem of the final good producer is:

$$\min_{\left(\tilde{y}_{cjt}, \{y_{ijt}\}_{i=1}^{N_{jt}} \, : \, j \in [0,1]\right)} \left\{ \int_0^1 \left(\tilde{p}_{cjt} \tilde{y}_{cjt} + \sum_{i=1}^{N_{jt}} p_{ijt} y_{ijt} \right) dj \right\}$$

$$\text{s.t. } Y_t = \exp\left\{ \int_0^1 \left(\frac{\gamma}{\gamma - 1} \right) \ln \left[\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}} + \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} y_{ijt}^{\frac{\eta-1}{\eta}} \right)^{\frac{(\gamma-1)\eta}{\gamma(\eta-1)}} \right] dj \right\}$$

The optimality conditions with respect to a superstar firm i and to the fringe, both belonging to industry j, yield the following inverse demand functions:

$$p_{ijt} = \hat{\omega}_{ijt} y_{ijt}^{-\frac{1}{\eta}} \tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} y_{jt}^{\frac{1}{\gamma}-1} Y_t \tag{C.2}$$

$$\tilde{p}_{cjt} = \tilde{y}_{cjt}^{-\frac{1}{\eta}} y_{jt}^{\frac{1}{\eta}-1} Y_t \tag{C.3}$$

respectively, recalling that $y_{jt} = \left(\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}} + \tilde{y}_{sjt}^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}}$, with

$$\tilde{y}_{cjt} = \int_{F_{jt}} y_{ckjt} dk \qquad \text{and} \qquad \tilde{y}_{sjt} = \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} y_{ijt}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}$$

It is easy to show that the inverse demand schedule above can be written in terms of prices by means of the appropriate price indices $(\tilde{p}_{cjt}, \tilde{p}_{sjt}, p_{jt})$, as done in the main text. In particular, we define $p_{jt} \equiv \left(\tilde{p}_{sjt}^{1-\gamma} + \tilde{p}_{cjt}^{1-\gamma} \right)^{\frac{1}{1-\gamma}}$, with $\tilde{p}_{cjt} \equiv w_t / q_{cjt}$ (as the fringe prices at marginal cost) and $\tilde{p}_{sjt} \equiv \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt}^{\eta} p_{ijt}^{1-\eta} \right)^{\frac{1}{1-\eta}}$.

Superstar's Problem Taking these demand schedules as given, the static problem of an individual superstar i in industry j consists of simultaneously choosing output y_{ijt} and advertising efforts ω_{ijt} to maximize static profits, taking the output and advertising choices of all other firms in the industry, i.e., $(\tilde{y}_{cjt}, \{y_{hjt}\}_{h\neq i})$ and $\{\omega_{hjt}\}_{h\neq i}$, as given. That is, superstar firm i solves:

$$\max_{y_{ijt},\omega_{ijt}} \left\{ \left(p_{ijt} - \frac{w_t}{q_{ijt}} \right) y_{ijt} - \chi_a \omega_{ijt}^{\phi_a} Y_t \right\} \text{ s.t. } p_{ijt} = \hat{\omega}_{ijt} y_{ijt}^{-\frac{1}{\eta}} \tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} y_{jt}^{\frac{1}{\gamma}-1} Y_t$$

where

$$\hat{\omega}_{ijt} \equiv \frac{1+\omega_{ijt}}{\frac{1}{N_{jt}}\sum_{k=1}(1+\omega_{kjt})}$$

The first-order conditions are, respectively:

$$\frac{\partial p_{ijt}}{\partial y_{ijt}} y_{ijt} + p_{ijt} = \frac{w_t}{q_{ijt}} \tag{C.4a}$$

$$\frac{\partial p_{ijt}}{\partial \omega_{ijt}} y_{ijt} = \chi_a \phi_a \omega_{ijt}^{\phi_a - 1} Y_t \tag{C.4b}$$

Output Choices, Market Shares and Markups Let us work out the first condition, equation (C.4a). Using the inverse demand, note that:

$$\frac{\partial p_{ijt}}{\partial y_{ijt}} = \hat{\omega}_{ijt} \Bigg\{ -\frac{1}{\eta} \underbrace{y_{ijt}^{-\frac{1}{\eta}-1} \tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} y_{jt}^{\frac{1}{\gamma}-1} Y_t}_{=\frac{1}{\hat{\omega}_{ijt}}\frac{p_{ijt}}{y_{ijt}}} + y_{ijt}^{-\frac{1}{\eta}} \Bigg[\left(\frac{\gamma-\eta}{\eta\gamma} \right) \tilde{y}_{sjt}^{\frac{1}{\eta}-1} \underbrace{\tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} \hat{\omega}_{ijt} y_{ijt}^{-\frac{1}{\eta}} y_{jt}^{\frac{1-\gamma}{\gamma}} Y_t}_{=p_{ijt}} - \left(\frac{\gamma-1}{\gamma} \right) \tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} y_{jt}^{\frac{1}{\gamma}-1} \underbrace{y_{jt}^{\frac{1}{\gamma}-1} \tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} \hat{\omega}_{ijt} y_{ijt}^{-\frac{1}{\eta}} Y_t}_{=p_{ijt}} \Bigg] \Bigg\}$$

Therefore:

$$\frac{\partial p_{ijt}}{\partial y_{ijt}} = -\frac{1}{\eta} \frac{p_{ijt}}{y_{ijt}} + \underbrace{\hat{\omega}_{ijt} y_{ijt}^{-\frac{1}{\eta}} \tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} y_{jt}^{\frac{1}{\gamma}-1}}_{=p_{ijt}/Y_t} \left[\left(\frac{\gamma-\eta}{\eta\gamma} \right) p_{ijt} \left(\frac{y_{jt}}{\tilde{y}_{sjt}} \right)^{\frac{\gamma-1}{\gamma}} - \left(\frac{\gamma-1}{\gamma} \right) p_{ijt} \right]$$

Using $\frac{\partial p_{ijt}}{\partial y_{ijt}} y_{ijt} + p_{ijt} = \frac{w_t}{q_{ijt}}$ by equation (C.4a) gives us a formula for the inverse markup:

$$\frac{1}{M_{ijt}} \equiv \frac{w_t/q_{ijt}}{p_{ijt}} = \left(\frac{\eta-1}{\eta}\right) - \frac{p_{ijt}y_{ijt}}{Y_t}\left[\left(\frac{\eta-\gamma}{\eta\gamma}\right)\left(\frac{y_{jt}}{\tilde{y}_{sjt}}\right)^{\frac{\gamma-1}{\gamma}} + \left(\frac{\gamma-1}{\gamma}\right)\right]$$

As every industry j gets the same share of output, we have $p_{jt}y_{jt} = Y_t$ (recall that the final good is the numeraire, $P_t = 1$). Therefore, we may define the market share of a leader i within its industry j (i.e., including the fringe) at time t as:

$$\sigma_{ijt} \equiv \frac{p_{ijt}y_{ijt}}{p_{jt}y_{jt}} = \frac{p_{ijt}y_{ijt}}{Y_t}$$

i.e., $\sigma_{ijt} = \hat{\omega}_{ijt} y_{ijt}^{1-\frac{1}{\eta}} \tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} y_{jt}^{\frac{1}{\gamma}-1}$. This allows us to write the inverse markup defined above as:

$$\frac{1}{M_{ijt}} = \left(\frac{\eta-1}{\eta}\right) - \sigma_{ijt}\left[\left(\frac{\eta-\gamma}{\eta\gamma}\right)\left(\frac{y_{jt}}{\tilde{y}_{sjt}}\right)^{\frac{\gamma-1}{\gamma}} + \left(\frac{\gamma-1}{\gamma}\right)\right] \tag{C.5}$$

The markup depends on two endogenous objects: $\left(\frac{y_{jt}}{\tilde{y}_{sjt}}\right)^{\frac{\gamma-1}{\gamma}}$ and σ_{ijt}. To make progress, note that both of these can be written in terms of relative outputs. To show this, first note that:

$$\left(\frac{y_{jt}}{\tilde{y}_{sjt}}\right)^{\frac{\gamma-1}{\gamma}} = \frac{\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}} + \tilde{y}_{sjt}^{\frac{\gamma-1}{\gamma}}}{\tilde{y}_{sjt}^{\frac{\gamma-1}{\gamma}}}\left(\frac{y_{ijt}}{y_{ijt}}\right)^{\frac{\gamma-1}{\gamma}} = \frac{\left(\frac{\tilde{y}_{cjt}}{y_{ijt}}\right)^{\frac{\gamma-1}{\gamma}} + \left(\frac{\tilde{y}_{sjt}}{y_{ijt}}\right)^{\frac{\gamma-1}{\gamma}}}{\left(\frac{\tilde{y}_{sjt}}{y_{ijt}}\right)^{\frac{\gamma-1}{\gamma}}}$$

where

$$\left(\frac{\tilde{y}_{sjt}}{y_{ijt}}\right)^{\frac{\gamma-1}{\gamma}} = \left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt}\left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}\frac{\gamma-1}{\gamma}}$$

Second, note that $\sigma_{ijt} = \hat{\omega}_{ijt}\left(\frac{y_{ijt}}{y_{jt}}\right)^{\frac{\eta-1}{\eta}}\left(\frac{y_{jt}}{\tilde{y}_{sjt}}\right)^{\frac{\eta-\gamma}{\eta\gamma}}$. Developing this expression:

$$\begin{aligned}\sigma_{ijt} = \hat{\omega}_{ijt}\left(\frac{y_{ijt}}{y_{jt}}\right)^{\frac{\eta-1}{\eta}}\left(\frac{y_{jt}}{\tilde{y}_{sjt}}\right)^{\frac{\eta-\gamma}{\eta\gamma}} &= \hat{\omega}_{ijt} y_{ijt}^{\frac{\eta-1}{\eta}} \frac{\left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} y_{hjt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}} + \left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} y_{hjt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}} \\ &= \hat{\omega}_{ijt}\frac{\left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} y_{hjt}^{\frac{\eta-1}{\eta}} y_{ijt}^{\frac{\eta-1}{\eta}\frac{\gamma(\eta-1)}{\gamma-\eta}}\right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}} + \left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} y_{hjt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}}\frac{y_{ijt}^{\frac{\gamma-1}{\gamma}}}{y_{ijt}^{\frac{\gamma-1}{\gamma}}}\end{aligned}$$

$$= \hat{\omega}_{ijt} \frac{\left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{\left(\frac{\tilde{y}_{cjt}}{y_{ijt}}\right)^{\frac{\gamma-1}{\gamma}} + \left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}}$$

From the last equation, note that:

$$\sigma_{ijt} \left(\frac{y_{jt}}{\tilde{y}_{sjt}}\right)^{\frac{\gamma-1}{\gamma}} = \hat{\omega}_{ijt} \frac{\left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{\left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}} = \frac{\hat{\omega}_{ijt}}{\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}} = \frac{p_{ijt} y_{ijt}}{\sum_{h=1}^{N_{jt}} p_{hjt} y_{hjt}} \equiv \tilde{\sigma}_{ijt}$$

implying that $\tilde{\sigma}_{ijt} = \hat{\omega}_{ijt} \left(\frac{y_{ijt}}{\tilde{y}_{sjt}}\right)^{\frac{\eta-1}{\eta}}$. Here, $\tilde{\sigma}_{ijt}$ denotes the market share of superstar i relative to other superstars within its industry (i.e., excluding the fringe). Plugging this definition back into equation (C.5), we have:

$$M_{ijt} = \left[\left(\frac{\eta-1}{\eta}\right) - \left(\frac{\gamma-1}{\gamma}\right)\sigma_{ijt} - \left(\frac{\eta-\gamma}{\eta\gamma}\right)\tilde{\sigma}_{ijt}\right]^{-1} \tag{C.6}$$

This is the expression for the markup written in the main text (equation (12)), where:

$$\sigma_{ijt} \equiv \frac{p_{ijt} y_{ijt}}{\tilde{p}_{cjt}\tilde{y}_{cjt} + \sum_{h=1}^{N_{jt}} p_{hjt} y_{hjt}} = \frac{\hat{\omega}_{ijt} \left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{\left(\frac{\tilde{y}_{cjt}}{y_{ijt}}\right)^{\frac{\gamma-1}{\gamma}} + \left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}} \tag{C.7}$$

$$\tilde{\sigma}_{ijt} \equiv \frac{p_{ijt} y_{ijt}}{\sum_{h=1}^{N_{jt}} p_{hjt} y_{hjt}} = \frac{\hat{\omega}_{ijt}}{\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}} \tag{C.8}$$

Using the inverse demand function, the relative output between two superstar firms i and k, and between some superstar firm i and the fringe, can be written as:

$$\left(\frac{y_{ijt}}{y_{kjt}}\right)^{\frac{1}{\eta}} = \frac{q_{ijt}}{q_{kjt}} \frac{\hat{\omega}_{ijt}}{\hat{\omega}_{kjt}} \frac{M_{kjt}}{M_{ijt}} \quad \text{and} \quad \frac{y_{ijt}}{\tilde{y}_{cjt}} = \frac{q_{ijt}}{q_{cjt}} \frac{\sigma_{ijt}}{\sigma_{cjt}} \frac{1}{M_{ijt}}$$

respectively, where $\sigma_{cjt} \equiv 1 - \sum_{h=1}^{N_{jt}} \sigma_{hjt}$. This shows that all that is needed to describe the static equilibrium conditions related to output, markups, and market shares, are the relative intrinsic qualities, which satisfy $\frac{q_{ijt}}{q_{kjt}} = (1+\lambda)^{n_{ijt}^k}$. Thus, the state $\{n_{ijt}^k\}_{i,k}$ is sufficient to describe the within-industry static allocation.

Advertising Choice Next, we show that the advertising choice also exhibits this sufficient-statistic property. For this, let us work out the optimality condition for advertising effort, equation (C.4b). First, we have:

$$\begin{aligned}
\frac{\partial p_{ijt}}{\partial \omega_{ijt}} &= \frac{p_{ijt}}{\hat{\omega}_{ijt}}\frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} + \underbrace{\hat{\omega}_{ijt} y_{ijt}^{-\frac{1}{\eta}} y_{jt}^{\frac{1}{\gamma}-1} \tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} Y_t}_{=p_{ijt}} \left(\sum_{h=1}^{N_{jt}} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} y_{hjt}^{\frac{\eta-1}{\eta}}\right) \left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right) \tilde{y}_{sjt}^{\frac{1}{\eta}-1} - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right) \tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} y_{jt}^{\frac{1}{\gamma}-1}\right] \\
&= \frac{p_{ijt}}{\hat{\omega}_{ijt}}\frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} + p_{ijt} \left(\sum_{h=1}^{N_{jt}} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right) \left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right) \underbrace{\tilde{y}_{sjt}^{\frac{1}{\eta}-1} y_{ijt}^{1-\frac{1}{\eta}}}_{=\tilde{\sigma}_{ijt}/\hat{\omega}_{ijt}} - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right) \underbrace{\tilde{y}_{sjt}^{\frac{1}{\eta}-\frac{1}{\gamma}} y_{jt}^{\frac{1}{\gamma}-1} y_{ijt}^{1-\frac{1}{\eta}}}_{=\sigma_{ijt}/\hat{\omega}_{ijt}}\right] \\
&= \frac{p_{ijt}}{\hat{\omega}_{ijt}} \left\{\frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} + \left(\sum_{h=1}^{N_{jt}} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right) \left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right) \tilde{\sigma}_{ijt} - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right) \sigma_{ijt}\right]\right\}
\end{aligned}$$

where, to arrive at the last line, we have used $p_{ijt} y_{ijt} = \sigma_{ijt} Y_t$ and $\tilde{\sigma}_{ijt} = \sigma_{ijt} \left(\frac{y_{jt}}{\tilde{y}_{sjt}}\right)^{\frac{\gamma-1}{\gamma}}$. Using $\frac{\partial p_{ijt}}{\partial \omega_{ijt}} \frac{y_{ijt}}{Y_t} = \chi_a \phi_a \omega_{ijt}^{\phi_a - 1}$ by the optimality condition, we then have:

$$\chi_a \phi_a \omega_{ijt}^{\phi_a-1} = \frac{\sigma_{ijt}}{\hat{\omega}_{ijt}} \left\{\frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} + \left(\sum_{h=1}^{N_{jt}} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right) \left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right) \tilde{\sigma}_{ijt} - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right) \sigma_{ijt}\right]\right\} \tag{C.9}$$

Using the definitions for $\{\hat{\omega}_{hjt}\}_{h=1}^{N_{jt}}$, we have:

$$\begin{aligned}
\frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} &= \frac{\frac{1}{N_{jt}} \sum_{h=1}^{N_{jt}} (1+\omega_{hjt}) - (1+\omega_{ijt}) \frac{1}{N_{jt}}}{\left[\frac{1}{N_{jt}} \sum_{h=1}^{N_{jt}} (1+\omega_{hjt})\right]^2} \\
&= \frac{1+\omega_{ijt}}{\frac{1}{N_{jt}} \sum_{h=1}^{N_{jt}} (1+\omega_{hjt})} \left[\frac{\frac{1}{N_{jt}} \sum_{h=1}^{N_{jt}} (1+\omega_{hjt})}{(1+\omega_{ijt}) \left(\frac{1}{N_{jt}} \sum_{h=1}^{N_{jt}} (1+\omega_{hjt})\right)} - \frac{\frac{1}{N_{jt}}}{\frac{1}{N_{jt}} \sum_{h=1}^{N_{jt}} (1+\omega_{hjt})}\right] \\
&= \hat{\omega}_{ijt} \left[\frac{1}{1+\omega_{ijt}} - \frac{1}{\sum_{h=1}^{N_{jt}} (1+\omega_{hjt})}\right] \\
&= \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \frac{\sum_{h \neq i} (1+\omega_{hjt})}{\sum_{h=1}^{N_{jt}} (1+\omega_{hjt})} \\
&= \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left(1 - \frac{\hat{\omega}_{ijt}}{N_{jt}}\right)
\end{aligned}$$

$$\forall h \neq i: \; \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} = \frac{-(1+\omega_{hjt})\frac{1}{N_{jt}}}{\left[\frac{1}{N_{jt}}\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})\right]^2} = -\frac{\hat{\omega}_{hjt}}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})}$$

Therefore, using equation (C.8):

$$\begin{aligned}
\sum_{h=1}^{N_{jt}} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}} &= \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \frac{\sum_{h\neq i}(1+\omega_{hjt})}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} - \sum_{h\neq i} \frac{\hat{\omega}_{hjt}}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}} \\
&= \frac{1}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \left[\frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \sum_{h\neq i}(1+\omega_{hjt}) - \sum_{h\neq i} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right] \\
&= \frac{1}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \left[\hat{\omega}_{ijt} \left(1 + \frac{1}{1+\omega_{ijt}} \sum_{h\neq i}(1+\omega_{hjt})\right) - \underbrace{\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(\frac{y_{kjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}}_{=\hat{\omega}_{ijt}/\tilde{\sigma}_{ijt}}\right] \\
&= \frac{1}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \left[\frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \sum_{k=1}^{N_{jt}}(1+\omega_{kjt}) - \frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right] \\
&= \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left(1 - \frac{1}{N_{jt}} \frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right)
\end{aligned}$$

Back into equation (C.9), we obtain:

$$\begin{aligned}
\chi_a \phi_a \omega_{ijt}^{\phi_a - 1} &= \frac{\sigma_{ijt}}{\hat{\omega}_{ijt}} \left\{ \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left(1 - \frac{\hat{\omega}_{ijt}}{N_{jt}}\right) \right. \\
&\qquad \left. + \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left(1 - \frac{1}{N_{jt}} \frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right) \left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right) \tilde{\sigma}_{ijt} - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right) \sigma_{ijt}\right] \right\} \\
&= \frac{\sigma_{ijt}}{1+\omega_{ijt}} \left\{ 1 - \frac{\hat{\omega}_{ijt}}{N_{jt}} + \left(1 - \frac{1}{N_{jt}} \frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right) \left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right) \tilde{\sigma}_{ijt} - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right) \sigma_{ijt}\right] \right\}
\end{aligned} \tag{C.10}$$

This is the expression for optimal advertising in the main text (equation (15)).

C.3 Derivation of the Growth Rate

This section derives the growth rate of the economy. Using the production function at the aggregate and industry levels, we can write:

$$
\begin{aligned}
\ln(Y_t) &= \int_0^1 \frac{\gamma}{\gamma-1} \ln\left[\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}} + \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} y_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right] dj \\
&= \int_0^1 \frac{\gamma}{\gamma-1} \ln\left[\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}} \left(1 + \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} \left(\frac{y_{ijt}}{\tilde{y}_{cjt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right)\right] dj \\
&= \int_0^1 \left[\ln\left(\tilde{y}_{cjt}\right) + \frac{\gamma}{\gamma-1} \ln\left(1 + \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} \left(\frac{y_{ijt}}{\tilde{y}_{cjt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right)\right] dj \\
&= \int_0^1 \left[\ln\left(\frac{q_{cjt}}{w_t^{rel}}\sigma_{cjt}\right) + \frac{\gamma}{\gamma-1} \ln\left(1 + \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} \left(\frac{y_{ijt}}{\tilde{y}_{cjt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right)\right] dj \\
&= \int_0^1 \left[\ln\left(\frac{q_{cjt}}{w_t^{rel}}\right) + \frac{1}{\gamma-1} \ln\left(1 + \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} \left(\frac{y_{ijt}}{\tilde{y}_{cjt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right)\right] dj \\
&= \int_0^1 \ln\left(\frac{q_{cjt}}{w_t^{rel}}\right) dj + \sum_{\Theta} f_t(\Theta)\mu_t(\Theta) \qquad \text{(C.11)}
\end{aligned}
$$

where $f_t(\Theta)$ is defined as

$$
f_t(\Theta) \equiv \frac{1}{\gamma-1} \ln\left(1 + \left(\sum_{i=1}^{N_t(\Theta)} \hat{\omega}_{it}(\Theta) \left(\frac{y_{it}}{\tilde{y}_{ct}}(\Theta)\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right) \qquad \text{(C.12)}
$$

and we have used that $\tilde{y}_{cjt} = q_{cjt} l_{cjt} = q_{cjt} \frac{\sigma_{cjt}}{w_t^{rel}}$ and

$$
\sigma_{cjt} = \left[1 + \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} \left(\frac{y_{ijt}}{\tilde{y}_{cjt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right]^{-1}
$$

to arrive at the final expression. For a time step of size $\Delta t \approx 0$, we have:

$$
\begin{aligned}
\ln(Y_{t+\Delta t}) - \ln(Y_t) &= -\ln(w_{t+\Delta t}^{rel}) + \ln(w_t^{rel}) + \ln(1+\lambda) \sum_{\Theta} p_t^{leader}(\Theta) \Delta t \mu_t(\Theta) \\
&\quad + \sum_{\Theta}\sum_{\Theta'} \left[f_t(\Theta') - f_t(\Theta)\right] p_t(\Theta,\Theta')\mu_t(\Theta)\Delta t + o(\Delta t) \qquad \text{(C.13)}
\end{aligned}
$$

Dividing through by Δt and taking the limit as $\Delta t \to 0$ we obtain:

$$g_t = -g_{w^{rel},t} + \ln(1+\lambda)\sum_{\Theta} p_t^{leader}(\Theta)\mu_t(\Theta) + \sum_{\Theta}\sum_{\Theta'}\left[f_t(\Theta') - f_t(\Theta)\right] p_t(\Theta,\Theta')\mu_t(\Theta) \quad \text{(C.14)}$$

Next, let $\hat{\Theta}$ denote the set of all industry states Θ. Let $h : \hat{\Theta} \to \mathbb{R}$ be a function. Then, in a stationary equilibrium:

$$\begin{aligned}
\mathbb{E}\left[\sum_{\Theta'} p(\Theta,\Theta')(h(\Theta') - h(\Theta))\right] &= \sum_{\Theta}\sum_{\Theta'} p(\Theta,\Theta')(h(\Theta') - h(\Theta))\mu(\Theta) \\
&= \sum_{\Theta}\sum_{\Theta'} p(\Theta,\Theta')h(\Theta')\mu(\Theta) - \sum_{\Theta}\sum_{\Theta'} p(\Theta,\Theta')h(\Theta)\mu(\Theta) \\
&= \sum_{\Theta'} h(\Theta')\sum_{\Theta} p(\Theta,\Theta')\mu(\Theta) - \sum_{\Theta} h(\Theta)\sum_{\Theta'} p(\Theta,\Theta')\mu(\Theta) \\
&= \sum_{\Theta'} h(\Theta')\mu(\Theta') - \sum_{\Theta} h(\Theta)\mu(\Theta) \\
&= \mathbb{E}\left[h(\Theta')\right] - \mathbb{E}\left[h(\Theta)\right] \\
&= 0
\end{aligned}$$

Combined with the fact that, in a balanced growth path, the relative wage is constant ($g_{w^{rel},t} = 0$) and there is a time-invariant distribution over industry states, $\mu_t(\Theta) = \mu(\Theta)$, we can calculate the BGP growth rate as

$$g = \ln(1+\lambda)\left(\sum_{\Theta} p^{leader}(\Theta)\mu(\Theta)\right) \quad \text{(C.15)}$$

C.4 Calculating Social Welfare Metrics

In this Appendix, we detail how to compute welfare for the representative household, as well as our measure of consumption-equivalent welfare changes, in a BGP. Along a BGP, household consumption grows at the same rate as aggregate output. Therefore, the stream of present-discounted value of utility from consumption can be summarized by two variables: an initial level of consumption, C_0, and the growth rate of the economy, g.

To compute the initial output, use equation (C.11) to write:

$$Y_0 = \exp\left(\int_0^1 \ln(q_{cj0})dj - \ln(w^{rel}) + \sum_{\Theta} f(\Theta)\mu(\Theta)\right) \quad \text{(C.16)}$$

In a BGP, all the terms are time-invariant, and we fix the average log productivity level of fringe firms at time zero, $\int_0^1 \ln(q_{cj0})dj$, to zero in all counterfactual economies without loss of generality.[10] The initial level of consumption is then given by:

$$C_0 = Y_0 \cdot \frac{C_0}{Y_0} = Y_0 \left(1 - \int_0^1 \sum_{i=1}^{N_{j0}} \chi z_{ij0}^{\phi} \, dj - \int_0^1 \sum_{i=1}^{N_{j0}} \chi_a \omega_{ij0}^{\phi_a} \, dj - \int_0^1 m_0 \nu X_{kj0}^{\epsilon} \, dj - \psi e_0^2 \right) \quad \text{(C.17)}$$

where we have used the aggregate resource constraint (equation (C.1)) at $t = 0$ on the right-hand side. The welfare of the representative household can be found by imposing BGP to equation (3):

$$W = \int_0^{+\infty} e^{-\rho t} \ln(C_t) \, dt = \frac{1}{\rho} \left(\ln(C_0) + \frac{g}{\rho} \right) \quad \text{(C.18)}$$

Using formulas (C.16) and (C.17), equation (24) readily follows. Finally, to compute consumption-equivalent welfare changes between two economies A and B in their BGPs, we compute the percentage change ς in lifetime consumption that the representative household of economy A would require to remain indifferent between living in economy A and living in economy B, that is:

$$W^B = \frac{1}{\rho} \left(\ln \left(C_0^A (1 + \varsigma) \right) + \frac{g^A}{\rho} \right) \quad \text{(C.19)}$$

Solving for ς, we get:

$$\varsigma = \frac{C_0^B}{C_0^A} \exp \left(\frac{g^B - g^A}{\rho} \right) - 1 \quad \text{(C.20)}$$

D Robustness Checks

D.1 Ex-Ante versus Ex-Post Preferences and Deceptive Advertising

One potential concern highlighted in the literature when evaluating the welfare effect of persuasive (taste-shifting) advertising relates to whether welfare should be computed using ex-ante or ex-post preferences (for discussions on this, see Dixit and Norman (1978) or Benhabib and Bisin (2011), among many others). In their book titled *Phishing for Phools*, Akerlof and Shiller (2015) highlight that companies often "exploit our psychological weaknesses and our ignorance through manipulation and deception". One example is deceptive advertising, which persuades consumers to buy certain products over others, but ex-post, the consumers find out that the products do not deliver what they imagined they would. This problem can be particularly severe for industries providing experience goods – products the quality of which can be accurately evaluated only after purchasing

[10] Because fringe firms keep a constant distance ζ with respect to their industry's leader by assumption, this assumption means that we keep the initial frontier technology level fixed across all counterfactual economies.

and experiencing them, such as books, movies, restaurants, and so on. A consumer can purchase a ticket to a widely advertised movie, only to find out that it does not "live up to the hype", and feel buyer's remorse in retrospect. On the other hand, in equilibrium, a consumer might miss out on purchasing under-advertised products that they would have enjoyed more, missing out on "hidden gems".

In our baseline experiments, we assume ex-ante and ex-post preferences coincide to evaluate the welfare implications of advertising, i.e., we evaluate welfare assuming that advertising influences consumers' welfare in the same way that it influences the preferences revealed by their demand. At the other extreme, one could argue that advertising is purely deceptive and that, as a result, welfare should be evaluated without any effect of advertising, i.e., $\hat{\omega}_{ijt} = 1$ for all $i, j,$ and t. The choice of which approach to follow is of course not neutral in terms of welfare implications of advertising.

Following this discussion, we propose an extension of our model in which we allow for deceptive advertising. We assume that, at every instant, advertising in any industry turns out to be (unexpectedly) purely manipulative with probability $\delta \in [0, 1]$.[11] The case with $\delta = 0$ corresponds to our baseline model where ex-ante and ex-post preferences coincide, whereas $\delta = 1$ implies that advertising is fully deceptive and does not lead to changes in preference shifters ex-post. Consequently, δ parametrizes how severe the deceptive advertising problem is in the overall economy. Note that none of the positive implications regarding the competitive equilibrium change, since purchases are still made according to the demand shifters $\hat{\omega}$ as in the baseline model. Only the (normative) welfare calculation is altered. However, when computing welfare, we now assume that with probability δ, advertising does not lead to a change in taste shifters, $\hat{\omega}_{it}$, and therefore they equal unity.

D.1.1 Derivations

We compute aggregate output when the degree of deception is δ, denoted $Y_t(\delta)$, as:

$$\ln(Y_t(\delta)) = \int_0^1 \ln\left(\frac{q_{cjt}}{w_t^{rel}}\right) dj + \sum_{\Theta} \Big(\delta h_t(\Theta) + (1-\delta) f_t(\Theta)\Big)\mu_t(\Theta) \tag{D.1}$$

where $f_t(\Theta)$ is defined in equation (C.12), and $h_t(\Theta)$ is the analogue of $f_t(\Theta)$ when advertising is deceptive, that is:

[11] In such instances, products of firms with $\hat{\omega} > 1$ are revealed to be "overhyped", and those with $\hat{\omega} < 1$ are revealed to be "hidden gems". In retrospect, the consumers would have preferred to purchase less of the prior and more of the latter, but their purchases are already made.

$$h_t(\Theta) \equiv \frac{1}{\gamma-1} \ln \left(1 + \left(\sum_{i=1}^{N_t(\Theta)} \left(\frac{y_{it}}{\tilde{y}_{ct}}(\Theta) \right)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}} \right)$$

Notice that measured output (equation (D.1)) coincides with output in the baseline model (equation (C.11)) when $\delta = 0$.

The remaining objects necessary to calculate welfare are computed as in the baseline model, as explained in Appendix C.4, so the consumption-equivalent welfare change between two economies A and B in their BGPs for a given level of deception δ is:

$$\varsigma(\delta) = \frac{C_0^B(\delta)}{C_0^A(\delta)} \exp\left(\frac{g^B - g^A}{\rho} \right) - 1$$

where $C_0(\delta)$ in both economies is computed as in equation (C.17), that is:

$$C_0(\delta) = Y_0(\delta) \left(1 - \int_0^1 \sum_{i=1}^{N_{j0}} \chi z_{ij0}^{\phi} \, dj - \int_0^1 \sum_{i=1}^{N_{j0}} \chi_a \omega_{ij0}^{\phi_a} \, dj - \int_0^1 m_0 \nu X_{kj0}^{\epsilon} \, dj - \psi e_0^2 \right)$$

D.1.2 Results

First, we repeat the advertising shutdown experiment for different values of $\delta \in \{0, 0.25, 0.50, 0.75, 1\}$. Table D.1 presents the associated consumption-equivalent welfare change numbers. As one might expect, as the deceptiveness of advertising δ is increased, the implied benefits of advertising diminish, as this reduces the static welfare gains from the consumers enjoying the cheaper-to-produce products of the leading superstars. In the extreme case scenario of purely deceptive advertising ($\delta = 1$), shutting down advertising is found to increase rather than decrease welfare.

TABLE D.1: DYNAMIC WELFARE IMPACT OF ADVERTISING SHUTDOWN WITH DECEPTIVE ADVERTISING

	$\delta = 0$	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1.00$
CEWC of Adv. Shutdown	-0.863%	-0.443%	-0.020%	0.404%	0.830%

Notes: This table presents the consumption-equivalent welfare change due to advertising shutdown in counterfactual economies where we assume that advertising in any industry is (unexpectedly) purely manipulative with probability δ.

Next, we calculate the welfare-maximizing advertising taxes and the associated positive changes in the economy under the extreme case of purely deceptive advertising ($\delta = 1$).

We find the optimal tax rate to be around 89.3%, and Table F.5 summarizes the changes to the economy. Interestingly, the optimal tax rate, while higher than in our baseline model with $\delta = 0$, is still below 100% – that is, it is not optimal to shut down advertising. The optimal tax increases welfare by 1.20% as opposed to the 0.83% gain from shutting down advertising altogether. This is because, even when we assume advertising to be completely deceptive, it still maintains the property of reducing static misallocation. Consequently, the static welfare gains from advertising are still positive, although lesser in magnitude compared to the baseline model. This allows the dynamic gains from shutting down advertising to dominate the static losses, flipping the welfare result as seen in Table D.1. However, a benevolent government would still choose to tax advertising at a high rate rather than to shut it down, so that the consumers can benefit from some improved static efficiency along with the dynamic gains.

Overall, this robustness check shows that as advertising gets closer to being purely deceptive, welfare losses from shutting down advertising decrease and can eventually turn into welfare gains, but the optimal tax rate is still below 100%, and there is still a role for advertising to fulfill, thanks to its property of alleviating static misallocation. Neither does the degree of the deceptiveness of advertising affect any of our positive (as opposed to normative) results on the effect of advertising on growth, markups, business dynamism, dynamic efficiency, and so on, independent of its influence on inferred welfare changes.

D.2 Non-Combative Advertising

In the baseline model, the shift in demand that results from an individual firm's expenditure in advertising is tampered by the advertising efforts of other firms: all else equal, an increase in a firm's advertising efforts will decrease the perceived quality of every other product in the industry. This makes advertising akin to a zero-sum game: if all firms were to choose the same advertising amount (ω), perceived quality ($\hat{\omega}$) would equal unity for all products, and consumers would receive no benefits despite all the resources spent on advertising. It would simply be wasteful spending.

However, the informative view on advertising highlights the fact that advertising can benefit consumers through making them aware of the existence of certain products, informing them of product characteristics, and helping them find the best product that matches their individual tastes.[12] Therefore, our combative advertising assumption in the baseline model might be too severe, and could be driving our results.

[12]For instance, Cavenaile *et al.* (2023) provide a microfoundation for the described mechanism, in which firms use advertising to expand the awareness sets of consumers over products, and help them achieve a better consumer-product match, increasing consumer welfare.

In this section, we extend the model by relaxing our baseline assumption regarding how perceived quality $\hat{\omega}$ is calculated, and generalize the degree of combativeness in the advertising technology. To this end, we assume that the perceived quality of variety i is now given by

$$\hat{\omega}_{ijt} \equiv \frac{1+\omega_{ijt}}{\Lambda + \frac{1-\Lambda}{N_{jt}}\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \tag{D.2}$$

where $\Lambda \in [0,1]$ is a parameter that governs the degree of advertising combativeness across firms. When $\Lambda = 0$, we return to our baseline model. When $\Lambda = 1$, we have $\hat{\omega}_{ijt} = 1 + \omega_{ijt}$. That is, the term in the denominator completely vanishes, and a firm's advertising does not directly affect the perceived quality of other products. As a consequence, if all firms chose the same advertising amount ω, the consumers would derive extra utility from the resources spent on advertising, which could be interpreted as capturing the informativeness of advertising in a reduced-form way.

D.2.1 Derivations

Under the specification in equation (25), the demand schedules faced by every firm are unchanged, so that equations (C.2)-(C.3) still hold. Likewise, the optimal firm-level markup is still given by equation (C.6). The advertising choice, however, is slightly different. While equation (C.9) continues to hold true, the set of derivatives $\{\frac{\partial\hat{\omega}_{hjt}}{\partial\omega_{ijt}}\}_{h=1}^{N_{jt}}$ is different. Using the definition of $\{\hat{\omega}_{hjt}\}_{h=1}^{N_{jt}}$ from (25), we have:

$$\begin{aligned}
\frac{\partial\hat{\omega}_{ijt}}{\partial\omega_{ijt}} &= \frac{\Lambda + \frac{1-\Lambda}{N_{jt}}\sum_{h=1}^{N_{jt}}(1+\omega_{hjt}) - (1+\omega_{ijt})\frac{1-\Lambda}{N_{jt}}}{\left[\Lambda + \frac{1-\Lambda}{N_{jt}}\sum_{h=1}^{N_{jt}}(1+\omega_{hjt})\right]^2} \\
&= \hat{\omega}_{ijt}\left[\frac{1}{1+\omega_{ijt}} - \frac{\frac{1-\Lambda}{N_{jt}}}{\Lambda + \frac{1-\Lambda}{N_{jt}}\sum_{h=1}^{N_{jt}}(1+\omega_{hjt})}\right] \\
&= \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}}\left[\frac{\Lambda + \frac{1-\Lambda}{N_{jt}}\sum_{h\neq i}(1+\omega_{hjt})}{\Lambda + \frac{1-\Lambda}{N_{jt}}\sum_{h=1}^{N_{jt}}(1+\omega_{hjt})}\right] \\
&= \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}}\left(1 - (1-\Lambda)\frac{\hat{\omega}_{ijt}}{N_{jt}}\right) \\
\forall h \neq i: \ \frac{\partial\hat{\omega}_{hjt}}{\partial\omega_{ijt}} &= \frac{-(1+\omega_{hjt})\frac{1-\Lambda}{N_{jt}}}{\left[\Lambda + \frac{1-\Lambda}{N_{jt}}\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})\right]^2} = -\frac{\hat{\omega}_{hjt}\frac{1-\Lambda}{N_{jt}}}{\Lambda + \frac{1-\Lambda}{N_{jt}}\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})}
\end{aligned}$$

Therefore:

$$\sum_{h=1}^{N_{jt}} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}} = \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left[\frac{\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{h\neq i}(1+\omega_{hjt})}{\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{h=1}^{N_{jt}}(1+\omega_{hjt})}\right] - \sum_{h\neq i} \frac{\hat{\omega}_{hjt}\frac{1-\Lambda}{N_{jt}}}{\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}$$

$$= \frac{1}{\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \left[\frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left(\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{h\neq i}(1+\omega_{hjt})\right) - \frac{1-\Lambda}{N_{jt}} \sum_{h\neq i} \hat{\omega}_{hjt} \left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right]$$

$$= \frac{1}{\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \left[\hat{\omega}_{ijt} \left(\frac{1-\Lambda}{N_{jt}} + \frac{1}{1+\omega_{ijt}} \left(\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{h\neq i}(1+\omega_{hjt})\right)\right) - \frac{1-\Lambda}{N_{jt}} \underbrace{\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(\frac{y_{kjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}}_{=\hat{\omega}_{ijt}/\tilde{\sigma}_{ijt}}\right]$$

$$= \frac{1}{\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \left[\frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left(\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{k=1}^{N_{jt}}(1+\omega_{kjt})\right) - \frac{1-\Lambda}{N_{jt}} \frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right]$$

$$= \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left(1 - \frac{1-\Lambda}{N_{jt}} \frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right)$$

Plugging back into equation (C.9), we obtain:

$$\chi_a \phi_a \omega_{ijt}^{\phi_a - 1} = \frac{\sigma_{ijt}}{\hat{\omega}_{ijt}} \left\{ \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left(1 - (1-\Lambda)\frac{\hat{\omega}_{ijt}}{N_{jt}}\right) + \frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}} \left(1 - \frac{1-\Lambda}{N_{jt}} \frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right) \left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right) \tilde{\sigma}_{ijt} - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right) \sigma_{ijt}\right] \right\}$$

$$= \frac{\sigma_{ijt}}{1+\omega_{ijt}} \left\{ 1 - (1-\Lambda)\frac{\hat{\omega}_{ijt}}{N_{jt}} + \left(1 - \frac{1-\Lambda}{N_{jt}} \frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right) \left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right) \tilde{\sigma}_{ijt} - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right) \sigma_{ijt}\right] \right\}$$

Note that setting $\Lambda = 0$, we return to the optimality condition of the baseline model (equation (C.10)). For $\Lambda = 1$, we obtain:

$$\chi_a \phi_a \omega_{ijt}^{\phi_a - 1} = \frac{\sigma_{ijt}}{1+\omega_{ijt}} \left[1 - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right) \sigma_{ijt} - \left(\frac{\eta-\gamma}{\gamma(\eta-1)}\right) \tilde{\sigma}_{ijt}\right]$$

D.2.2 Results

Unlike the previous extension, the assumptions regarding Λ have positive implications for the economy as well as normative, and hence, the extended model needs to be re-calibrated. To prove the robustness of our results, we pick the extreme value of $\Lambda = 1$ as opposed to our baseline's $\Lambda = 0$, re-calibrate the model, and repeat the experiments. Table F.6 presents the calibrated parameter values and the details of the calibration procedure, whereas Table F.7 summarizes the results of the experiments.

As one might expect, the extended model with $\Lambda = 1$ makes advertising more useful from a social perspective, and therefore the welfare cost of shutting down advertising is now much higher at 5.38% compared to the 0.86% calculated using the baseline model. This large increase primarily owes to the 80% higher impact of the shutdown on initial output, due to the increased direct benefit of advertising on welfare. In both models, shutting down advertising affects all economic quantities of interest in the same direction, although exact magnitudes vary.

Moving on to the optimal taxation experiment, we find that it is still optimal to tax advertising rather than subsidize it. The optimal tax rate is now 28.6% compared to the 62.9% found in the baseline, which is lower, but still quite significant. Given that we considered the extreme case of $\Lambda = 1$, this result proves the robustness of our taxation result regarding advertising – although the degree of advertising combativeness Λ influences how high the optimal advertising tax should be, even if we make advertising completely non-combative, a benevolent government should still tax advertising rather than subsidize it. Overall, this extension demonstrates the robustness of our quantitative results in direction, if not in magnitude.

We have also performed another robustness check in which we have calibrated the value of Λ. A higher value of Λ lowers the direct effect of a firm's advertising on the perceived quality of other firms, reducing the degree of "combativeness". Based on this insight, we use the correlation between a firm's relative sales, its own advertising, and total advertising by its competitors in the data to discipline the value of Λ. In particular, we regress the firm's relative sales on its own log advertising, and the log total advertising by other firms in the same industry, using the same controls and normalization for advertising variables as in Table A.1. The regression results can be found in the first column of Table A.2.

We target the ratio of the regression coefficient on own advertising to the regression coefficient on competitors' advertising, i.e., β_3/β_4 from regression (2). The calibration suggests a value for Λ that is very close to zero (0.0199) as assumed in our baseline model (see Table F.8 for the calibration results). As a consequence, the results of our counterfactual experiments remain mostly unaffected quantitatively as can be seen in Table F.9. In particular, an advertising shutdown leads to an increase in the growth rate of the economy by 3.62% (compared to 3.26% in the baseline calibration) and a decrease in the level of initial output by 5.34% (compared to 4.63% in the baseline model), which result in a welfare loss of 1.32% in consumption equivalent terms (compared 0.86% to in the baseline calibration). The optimal tax rate is lower than in our baseline model at around 54%. This confirms the robustness of our results to the degree of advertising combativeness suggested by the data.

D.3 Bertrand Competition

In the baseline model, superstar firms compete in quantities in a static Cournot game. One may wonder whether our results are contingent upon this assumption. To alleviate such concerns, we solve our model with the alternative assumption of competition in prices *à la* Bertrand, calibrate it, and repeat the quantitative experiments. This reveals that almost all of our results are maintained.

D.3.1 Derivations

In the Bertrand-pricing version of the model, the static problem of an individual superstar i in industry j consists of simultaneously choosing price p_{ijt} and advertising efforts ω_{ijt} to maximize static profits, taking the prices and advertising choices of all other firms in the industry, $(\tilde{p}_{cjt}, \{p_{hjt}\}_{h\neq i})$ and $\{\omega_{hjt}\}_{h\neq i}$, as given. That is, superstar firm i solves:

$$\max_{p_{ijt},\omega_{ijt}} \left\{ \left(p_{ijt} - \frac{w_t}{q_{ijt}} \right) y_{ijt} - \chi_a \omega_{ijt}^{\phi_a} Y_t \right\} \text{ s.t. } y_{ijt} = \hat{\omega}_{ijt}^{\eta} p_{ijt}^{-\eta} \tilde{p}_{sjt}^{\eta-\gamma} p_{jt}^{\gamma-1} Y_t$$

where $\tilde{p}_{sjt} = \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt}^{\eta} p_{ijt}^{1-\eta} \right)^{\frac{1}{1-\eta}}$, $p_{jt} = \left(\tilde{p}_{sjt}^{1-\gamma} + \tilde{p}_{cjt}^{1-\gamma} \right)^{\frac{1}{1-\gamma}}$, and $\hat{\omega}_{ijt} = \frac{1+\omega_{ijt}}{N_{jt}^{-1} \sum_{k=1}(1+\omega_{kjt})}$. The first-order conditions for p_{ijt} and ω_{ijt} are, respectively:

$$y_{ijt} + \left(p_{ijt} - \frac{w_t}{q_{ijt}} \right) \frac{\partial y_{ijt}}{\partial p_{ijt}} = 0 \tag{D.3a}$$

$$\left(p_{ijt} - \frac{w_t}{q_{ijt}} \right) \frac{\partial y_{ijt}}{\partial \omega_{ijt}} = \chi_a \phi_a \omega_{ijt}^{\phi_a - 1} Y_t \tag{D.3b}$$

Let us first work out condition (D.3a). Using the demand function, note:

$$\frac{\partial y_{ijt}}{\partial p_{ijt}} = \hat{\omega}_{ijt} \Bigg\{ -\eta \underbrace{p_{ijt}^{-\eta-1} \tilde{p}_{sjt}^{\eta-\gamma} p_{jt}^{\gamma-1} Y_t}_{= \hat{\omega}_{ijt}^{-\eta} \frac{y_{ijt}}{p_{ijt}}} + p_{ijt}^{-\eta} \Bigg[(\eta - \gamma) \tilde{p}_{sjt}^{\eta-1} \underbrace{\hat{\omega}_{ijt}^{\eta} \tilde{p}_{sjt}^{\eta-\gamma} p_{ijt}^{-\eta} p_{jt}^{\gamma-1} Y_t}_{= y_{ijt}}$$

$$+ (\gamma - 1) \tilde{p}_{sjt}^{\eta-\gamma} p_{jt}^{\gamma-1} \underbrace{\hat{\omega}_{ijt}^{\eta} p_{ijt}^{-\eta} \tilde{p}_{sjt}^{\eta-\gamma} p_{jt}^{\gamma-1} Y_t}_{= y_{ijt}} \Bigg] \Bigg\}$$

Therefore:

$$\frac{\partial y_{ijt}}{\partial p_{ijt}} = -\eta \frac{y_{ijt}}{p_{ijt}} + \underbrace{\hat{\omega}_{ijt}^{\eta} p_{ijt}^{-\eta} \tilde{p}_{sjt}^{\eta-\gamma} p_{jt}^{\gamma-1}}_{=y_{ijt}/Y_t} \left[(\eta - \gamma) \left(\frac{\tilde{p}_{sjt}}{p_{jt}} \right)^{\gamma-1} y_{ijt} + (\gamma - 1) y_{ijt} \right] \tag{D.4}$$

Using $\frac{\partial p_{ijt}}{\partial y_{ijt}} y_{ijt} + p_{ijt} = \frac{w_t}{q_{ijt}}$ by equation (D.3a) gives us a formula for the markup:

$$M_{ijt} \equiv \frac{p_{ijt}}{w_t / q_{ijt}} = \frac{\mathcal{E}_{ijt}}{\mathcal{E}_{ijt} - 1}$$

where $\mathcal{E}_{ijt} \equiv -\frac{p_{ijt}}{y_{ijt}} \frac{\partial y_{ijt}}{\partial p_{ijt}}$ is the price-elasticity of demand. Using that $\sigma_{ijt} = \frac{p_{ijt} y_{ijt}}{Y_t}$ is the firm's market share, note from equation (D.4) that:

$$\mathcal{E}_{ijt} = -\frac{p_{ijt}}{y_{ijt}} \frac{\partial y_{ijt}}{\partial p_{ijt}} = \eta - \sigma_{ijt} \left[(\eta - \gamma) \left(\frac{\tilde{p}_{sjt}}{p_{jt}} \right)^{\gamma-1} + (\gamma - 1) \right] \tag{D.5}$$

Finally, it is easy to show that $\sigma_{ijt} \left(\frac{\tilde{p}_{sjt}}{p_{jt}} \right)^{\gamma-1} = \tilde{\sigma}_{ijt}$. Putting our results together, the Bertrand-equilibrium markup can be written as:

$$M_{ijt} = \frac{\eta - (\eta - \gamma)\tilde{\sigma}_{ijt} - (\gamma - 1)\sigma_{ijt}}{\eta - 1 - (\eta - \gamma)\tilde{\sigma}_{ijt} - (\gamma - 1)\sigma_{ijt}}$$

Next, we move to the optimality condition for advertising, equation (D.3b). We have:

$$\begin{aligned}
\frac{\partial y_{ijt}}{\partial \omega_{ijt}} &= \eta \frac{y_{ijt}}{\hat{\omega}_{ijt}} \frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} \\
&+ \underbrace{\hat{\omega}_{ijt}^{\eta} p_{ijt}^{-\eta} p_{jt}^{\gamma-1} \tilde{p}_{sjt}^{\eta-\gamma} Y_t}_{=y_{ijt}} \left(\sum_{h=1}^{N_{jt}} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} \left(\frac{p_{hjt}}{\hat{\omega}_{hjt}} \right)^{1-\eta} \right) \left[\left(\frac{\eta(\gamma - \eta)}{\eta - 1} \right) \tilde{p}_{sjt}^{\eta-1} - \left(\frac{\eta(\gamma - 1)}{\eta - 1} \right) \tilde{p}_{sjt}^{\eta-\gamma} p_{jt}^{\gamma-1} \right] \\
&= \eta \frac{y_{ijt}}{\hat{\omega}_{ijt}} \frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} + y_{ijt} \left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt}^{\eta-1} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} \left(\frac{p_{hjt}}{p_{ijt}} \right)^{1-\eta} \right) \left[\left(\frac{\eta(\gamma - \eta)}{\eta - 1} \right) \underbrace{\left(\frac{\tilde{p}_{sjt}}{p_{ijt}} \right)^{\eta-1}}_{=\tilde{\sigma}_{ijt}/\hat{\omega}_{ijt}^{\eta}} - \left(\frac{\eta(\gamma - 1)}{\eta - 1} \right) \underbrace{\tilde{p}_{sjt}^{\eta-\gamma} p_{jt}^{\gamma-1} p_{ijt}^{1-\eta}}_{=\sigma_{ijt}/\hat{\omega}_{ijt}^{\eta}} \right] \\
&= \frac{y_{ijt}}{\hat{\omega}_{ijt}^{\eta}} \left\{ \eta \hat{\omega}_{ijt}^{\eta-1} \frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} + \left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt}^{\eta-1} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} \left(\frac{p_{hjt}}{p_{ijt}} \right)^{1-\eta} \right) \left[\left(\frac{\eta(\gamma - \eta)}{\eta - 1} \right) \tilde{\sigma}_{ijt} - \left(\frac{\eta(\gamma - 1)}{\eta - 1} \right) \sigma_{ijt} \right] \right\}
\end{aligned}$$

where, to arrive to the last line, we have used $p_{ijt} y_{ijt} = \sigma_{ijt} Y_t$ and $\tilde{\sigma}_{ijt} = \sigma_{ijt} \left(\frac{\tilde{p}_{sjt}}{p_{jt}} \right)^{\gamma-1}$. Using $\frac{\partial y_{ijt}}{\partial \omega_{ijt}} \frac{p_{ijt}}{Y_t} (1 - M_{ijt}^{-1}) = \chi_a \phi_a \omega_{ijt}^{\phi_a - 1}$ by the optimality condition, we then have:

$$\chi_a \phi_a \omega_{ijt}^{\phi_a - 1} = \frac{\sigma_{ijt}(1 - M_{ijt}^{-1})}{\hat{\omega}_{ijt}^{\eta}} \left\{ \eta \hat{\omega}_{ijt}^{\eta-1} \frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} \right. \tag{D.6}$$

$$+\left(\sum_{h=1}^{N_{jt}}\hat{\omega}_{hjt}^{\eta-1}\frac{\partial\hat{\omega}_{hjt}}{\partial\omega_{ijt}}\left(\frac{p_{hjt}}{p_{ijt}}\right)^{1-\eta}\right)\left[\left(\frac{\eta(\gamma-\eta)}{\eta-1}\right)\tilde{\sigma}_{ijt}-\left(\frac{\eta(\gamma-1)}{\eta-1}\right)\sigma_{ijt}\right]\Bigg\}$$

Recall that $\frac{\partial\hat{\omega}_{ijt}}{\partial\omega_{ijt}}=\frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}}\left(1-\frac{\hat{\omega}_{ijt}}{N_{jt}}\right)$ and $\frac{\partial\hat{\omega}_{hjt}}{\partial\omega_{ijt}}=-\frac{\hat{\omega}_{hjt}}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})}$, $\forall h\neq i$. Then, we have:

$$\begin{aligned}
\sum_{h=1}^{N_{jt}}\hat{\omega}_{hjt}^{\eta-1}\frac{\partial\hat{\omega}_{hjt}}{\partial\omega_{ijt}}\left(\frac{p_{hjt}}{p_{ijt}}\right)^{1-\eta} &= \frac{\hat{\omega}_{ijt}^{\eta}}{1+\omega_{ijt}}\frac{\sum_{h\neq i}(1+\omega_{hjt})}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})}-\sum_{h\neq i}\frac{\hat{\omega}_{hjt}^{\eta}}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})}\left(\frac{p_{hjt}}{p_{ijt}}\right)^{1-\eta}\\
&=\frac{1}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})}\left[\frac{\hat{\omega}_{ijt}^{\eta}}{1+\omega_{ijt}}\sum_{h\neq i}(1+\omega_{hjt})-\sum_{h\neq i}\hat{\omega}_{hjt}^{\eta}\left(\frac{p_{hjt}}{p_{ijt}}\right)^{1-\eta}\right]\\
&=\frac{1}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})}\left[\hat{\omega}_{ijt}^{\eta}\left(1+\frac{1}{1+\omega_{ijt}}\sum_{h\neq i}(1+\omega_{hjt})\right)-\underbrace{\sum_{k=1}^{N_{jt}}\hat{\omega}_{kjt}^{\eta}\left(\frac{p_{kjt}}{p_{ijt}}\right)^{1-\eta}}_{=\hat{\omega}_{ijt}^{\eta}/\tilde{\sigma}_{ijt}}\right]\\
&=\frac{1}{\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})}\left[\frac{\hat{\omega}_{ijt}^{\eta}}{1+\omega_{ijt}}\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})-\frac{\hat{\omega}_{ijt}^{\eta}}{\tilde{\sigma}_{ijt}}\right]\\
&=\frac{\hat{\omega}_{ijt}^{\eta}}{1+\omega_{ijt}}\left(1-\frac{1}{N_{jt}}\frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right)
\end{aligned}$$

Substituting this back into equation (D.6), we obtain:

$$\begin{aligned}
\chi_a\phi_a\omega_{ijt}^{\phi_a-1} &= \frac{\sigma_{ijt}(1-M_{ijt}^{-1})}{\hat{\omega}_{ijt}^{\eta}}\Bigg\{\eta\frac{\hat{\omega}_{ijt}^{\eta}}{1+\omega_{ijt}}\left(1-\frac{\hat{\omega}_{ijt}}{N_{jt}}\right)\\
&\qquad+\frac{\hat{\omega}_{ijt}^{\eta}}{1+\omega_{ijt}}\left(1-\frac{1}{N_{jt}}\frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right)\left[\left(\frac{\eta(\gamma-\eta)}{\eta-1}\right)\tilde{\sigma}_{ijt}-\left(\frac{\eta(\gamma-1)}{\eta-1}\right)\sigma_{ijt}\right]\Bigg\}\\
&=\eta\frac{\sigma_{ijt}(1-M_{ijt}^{-1})}{1+\omega_{ijt}}\left\{1-\frac{\hat{\omega}_{ijt}}{N_{jt}}+\left(1-\frac{1}{N_{jt}}\frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right)\left[\left(\frac{\gamma-\eta}{\eta-1}\right)\tilde{\sigma}_{ijt}-\left(\frac{\gamma-1}{\eta-1}\right)\sigma_{ijt}\right]\right\}
\end{aligned}$$

D.3.2 Results

We calibrate the parameters of this alternative model using the same methodology as the baseline analysis. Table F.10 presents the calibrated parameter values and the model fit. Using this calibrated model, we repeat the counterfactual experiments, the results of which are displayed in Table F.11. The advertising shutdown experiment reveals once again that advertising and innovation are substitutes at the economy level. Shutting down advertising boosts innovation, business dynamism, economic growth, and the labor share as in the baseline, whereas the markups and their dispersion go down. Similarly, the shutdown adversely affects initial output, as it increases the static misallocation across superstars. From a normative perspective, it is found that the dynamic gains slightly dominate the static losses this time, leading to a minor gain in consumption-equivalent welfare similar to what we observed in the deceptive advertising extension under high values of δ.

However, as was the case in the model with deceptive advertising, advertising is still found to be socially valuable. Repeating the optimal advertising tax experiment reveals that taxing advertising heavily is still preferable to shutting it down altogether. The optimal linear tax rate is found to be 90.65%, and adopting this tax rate delivers a consumption-equivalent welfare change of 1.85%, which is more than double the gains from shutting down advertising.

To sum up, moving from competition in quantities to competition in prices and re-calibrating the model using the same methodology serves to reduce the average level of static misallocation across industries. When static misallocation is lower through assumption, so are the quantitative gains from reducing it via advertising. As in Burstein *et al.* (2020), we assume Cournot competition in our baseline analysis due to its ability to generate more variation in markups and more realistic market share distributions consistent with what is observed for large firms in the United States, but the fact remains that most of our results go through regardless of the specific assumption on whether firms compete in prices or quantities.

D.4 $\bar{N}$ and $\bar{n}$

Finally, we conduct robustness checks in which we separately increase $\bar{N}$, from 4 to 5, and $\bar{n}$, from 5 to 6. The results of the recalibration with $\bar{N} = 5$ are summarized in Table F.14. Results of the advertising shutdown experiment and for the optimal tax can be found in Table F.15. Results from this recalibration are very similar to the baseline calibration. An advertising shutdown causes growth to increase by 3.48% (3.26% in the baseline calibration) while the level of output decreases by 4.96% (4.63% in the baseline calibration), leading to an overall decrease in welfare by 1.06% (0.86% in the baseline calibration). The optimal tax rate is also similar at 57.6% compared to 62.9% in the baseline model. This confirms that our baseline results do not significantly depend on the assumed maximum number of superstars per industry.

In another robustness exercise, we have increased $\bar{n}$ from 5 to 6 and recalibrated the model. Once again, our main results remain unaffected. Results can be found in Table F.13. Shutting down advertising increases growth by 2.77% while decreasing the level of output by 4.67%, which leads to a decrease in welfare by 1.19%.

E Social Planner's Problem

There exist both static and dynamic distortions in the economy. Statically, there are efficiency losses from the misallocation of labor both within and across industries due to the presence of market power. Moreover, there are efficiency losses coming from the choice of advertising, as firms do not internalize the effects that their advertising choices have on markup dispersion and the profits of other firms. Dynamically, resources for R&D are misallocated because firms fail to internalize the positive aggregate effects of their innovations on economic growth, as well as the negative contribution of their innovation resulting from business-stealing externalities.

E.1 The Complete Social Planner's Problem

The goal of the social planner is to maximize the lifetime utility of the representative household subject to the technological constraints of the economy. Given the initial conditions, $\mu_0(\Theta), m_0$, and aggregate productivity Q_0, the full problem can be stated as follows:

$$\max_{\left[\left[\{l_{ijt},\omega_{ijt},z_{ijt}\}_{i=1}^{N_{jt}},[l_{ckjt},X_{kjt}]_{k=0}^{m_t}:j\in[0,1]\right],e_t:t\in\mathbb{R}_+\right]} \int_0^{+\infty} e^{-\rho t}\ln(C_t)dt \tag{E.1}$$

subject to

$$C_t + R_tY_t + A_tY_t \leq Y_t \tag{E.2a}$$

$$R_t = \int_0^1 \left(\sum_{i=1}^{N_{jt}} \chi z_{ijt}^{\phi} + \int \nu X_{kjt}^{\epsilon} dk\right) dj + \psi e_t^2 \tag{E.2b}$$

$$A_t = \int_0^1 \sum_{i=1}^{N_{jt}} \chi_a \omega_{ijt}^{\phi_a}\, dj \tag{E.2c}$$

$$\ln(Y_t) = \int_0^1 \ln(y_{jt})dj \tag{E.2d}$$

$$y_{jt} = \left(\tilde{y}_{sjt}^{\frac{\gamma-1}{\gamma}} + \tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}} \tag{E.2e}$$

$$\tilde{y}_{sjt} = \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} y_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \tag{E.2f}$$

$$\hat{\omega}_{ijt} = \frac{1+\omega_{ijt}}{\frac{1}{N_{jt}}\sum_{i=1}^{N_{jt}}(1+\omega_{ijt})} \tag{E.2g}$$

$$\tilde{y}_{cjt} = \int y_{ckjt}dk \tag{E.2h}$$

$$y_{ijt} = q_{ijt} l_{ijt} \tag{E.2i}$$

$$y_{ckjt} = q_{cjt} l_{ckjt} \tag{E.2j}$$

$$\int_0^1 \left(\sum_{i=1}^{N_{jt}} l_{ijt} + \int l_{ckjt} dk \right) dj \leq L = 1 \tag{E.2k}$$

$$q_{jt}^{leader} = \max\{q_{1jt}, ..., q_{N_{jt}jt}\} \tag{E.2l}$$

$$q_{cjt} = \zeta q_{jt}^{leader} \tag{E.2m}$$

$$\{q_{1jt}, ..., q_{N_{jt}jt}\} = \left\{ q_{jt}^{leader}, \frac{q_{jt}^{leader}}{(1+\lambda)^{\vec{n}_{jt}(1)}}, ..., \frac{q_{jt}^{leader}}{(1+\lambda)^{\vec{n}_{jt}(N_{jt}-1)}} \right\} \tag{E.2n}$$

$$\Theta_{jt} = (N_{jt}, \vec{n}_{jt}) \tag{E.2o}$$

$$Q_t = \int \ln(q_{jt}^{leader}) dj \tag{E.2p}$$

$$\frac{\dot{Q}_t}{Q_t} = \ln(1+\lambda) \sum_{\Theta} p_t^{leader}(\Theta) \mu_t(\Theta) \tag{E.2q}$$

$$\dot{\mu}_t(\Theta) = \sum_{\Theta'} p_t(\Theta', \Theta) \mu_t(\Theta') - \sum_{\Theta'} p_t(\Theta, \Theta') \mu_t(\Theta) \tag{E.2r}$$

$$\sum_{\Theta} \mu_t(\Theta) = 1 \tag{E.2s}$$

$$\dot{m}_t = e_t - \tau m_t \tag{E.2t}$$

where equation (E.2a) is the resource constraint; equation (E.2b) is total R&D and business creation investment as a share of GDP; equation (E.2c) is the advertising share of GDP; equations (E.2d) to (E.2j) define production technologies at different levels of aggregation; equation (E.2g) defines the advertising shifter of superstars; equation (E.2k) is the aggregate labor feasibility constraint; equation (E.2l) defines the productivity of the industry leader; equation (E.2m) defines the productivity of each small firm in the competitive fringe, equation (E.2n) defines the vector of productivity step sizes; equation (E.2o) defines the relevant state of an industry, which can be summarized by the number of superstars in the industry (N_{jt}) and the number of productivity steps between each firm and the industry leader $\vec{n}_{jt}$; equation (E.2p) defines the average (log) productivity of leaders across industries; equation (E.2q) defines the growth rate of average productivity, where $\mu_t(\Theta)$ is the mass of industries in state Θ and $p_t^{leader}(\Theta)$ is the arrival rate at which one of the industry leaders innovates; equation (E.2r) is the law of motion of the industry distribution; equation (E.2s) states that the mass of industries has to sum to one; and equation (E.2t) is the law of motion of the mass of small firms.

The social planner maximizes welfare by choosing an allocation of labor and advertising to every superstar firm i in industry j at time t (l_{ijt}, ω_{ijt}) and labor to every small firm k

in industry j at time t (l_{kcjt}). The social planner also chooses R&D innovation policies for every superstar firm (z_{ijt}) and small firm (X_{kjt}) as well as the entry policy of entrepreneurs (e_t). Since small firms within the fringe of a given industry are symmetric, we can write the total labor allocation to small firms in industry j at time t as $l_{cjt} = m_t l_{kcjt}$ and the Poisson rate of emergence of a new superstar as $X_{jt} = m_t X_{kjt}$.

Even though this is a large problem to solve, it can be split into a static problem and a dynamic problem. By monotonicity of preferences, the final good and labor feasibility constraints (equations (E.2a) and (E.2k)) must bind with equality. Therefore, for a given distribution of productivities $[\{q_{ijt}\}_{i=1}^{N_{jt}}, q_{cjt}]_{j=0}^{1}$, the social planner wants to maximize total output Y_t net of advertising costs $A_t Y_t$ for all t, subject to the production technologies and the labor feasibility constraint. We solve this static output maximization problem next.

E.2 Static Output Maximization

Given the productivity distribution $\left[\{q_{ijt}\}_{i=1}^{N_{jt}}, q_{cjt} : j \in [0,1]\right]$, the social planner's static problem at time t is:

$$\max_{\left[\{l_{ijt},\omega_{ijt}\}_{i=1}^{N_{jt}}, l_{cjt}: j\in[0,1]\right]} \left\{ \frac{\gamma}{\gamma-1} \int_0^1 \ln\left(\left(\sum_{i=1}^{N_{jt}} \frac{1+\omega_{ijt}}{\frac{1}{N_{jt}}\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})} \left(l_{ijt} q_{ijt}\right)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}} + \left(l_{cjt} q_{cjt}\right)^{\frac{\gamma-1}{\gamma}} \right) dj \right.$$
$$\left. + \ln\left(1 - \int_0^1 \chi_a \sum_{i=1}^{N_{jt}} \omega_{ijt}^{\phi_a}\, dj\right) \right\} \tag{E.3}$$

$$\text{such that} \int_0^1 \left(\sum_{i=1}^{N_{jt}} l_{ijt} + l_{cjt} \right) dj = 1 \tag{E.4}$$

The first order conditions with respect to the labor input choices are:

$$\frac{\left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(l_{kjt} q_{kjt}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{\left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(l_{kjt} q_{kjt}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}} + \left(l_{cjt} q_{cjt}\right)^{\frac{\gamma-1}{\gamma}}} \hat{\omega}_{ijt} q_{ijt}^{\frac{\eta-1}{\eta}} l_{ijt}^{-\frac{1}{\eta}} = \vartheta_t, \quad \forall i \in \{1, ..., N_{jt}\}, \forall j \in [0,1] \tag{E.5}$$

$$\frac{q_{cjt}^{\frac{\gamma-1}{\gamma}} l_{cjt}^{-\frac{1}{\gamma}}}{\left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(l_{kjt} q_{kjt}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}} + \left(l_{cjt} q_{cjt}\right)^{\frac{\gamma-1}{\gamma}}} = \vartheta_t, \quad \forall j \in [0,1] \tag{E.6}$$

where $\vartheta_t > 0$ is the Lagrange multiplier associated with the labor feasibility constraint (E.4), and recall that $\hat{\omega}_{ijt} = \frac{1+\omega_{ijt}}{\frac{1}{N_{jt}}\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})}$. From these equations, it follows that:

$$\vartheta_t \sum_{i=1}^{N_{jt}} l_{ijt} = \frac{\left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(l_{kjt} q_{kjt}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}}{\left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(l_{kjt} q_{kjt}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}} + \left(l_{cjt} q_{cjt}\right)^{\frac{\gamma-1}{\gamma}}} \tag{E.7}$$

$$\vartheta_t l_{cjt} = \frac{\left(q_{cjt} l_{cjt}\right)^{\frac{\eta-1}{\eta}}}{\left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(l_{kjt} q_{kjt}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}} + \left(l_{cjt} q_{cjt}\right)^{\frac{\gamma-1}{\gamma}}} \tag{E.8}$$

Therefore, $\vartheta_t \left(\sum_{i=1}^{N_{jt}} l_{ijt} + l_{cjt}\right) = 1$. As $\int_0^1 \left(\sum_{i=1}^{N_{jt}} l_{ijt} + l_{cjt}\right) dj = 1$, we have $\vartheta_t = 1$. Consequently,

$$\sum_{i=1}^{N_{jt}} l_{ijt} + l_{cjt} = 1, \quad \forall j \in [0,1] \tag{E.9}$$

meaning that the planner allocates equal labor to all industries. To find the within-industry allocation of labor, we use equations (E.5) and (E.6) to establish:

$$\frac{l_{ijt}}{l_{kjt}} = \left(\frac{\hat{\omega}_{ijt}}{\hat{\omega}_{kjt}}\right)^{\eta} \left(\frac{q_{ijt}}{q_{kjt}}\right)^{\eta-1}, \forall i,k \in \{1,...,N_{jt}\}, \forall j \in [0,1] \tag{E.10}$$

$$\frac{l_{ijt}}{l_{cjt}} = \hat{\omega}_{ijt}^{\frac{\eta(\gamma-1)}{\eta-1}} \left(\frac{q_{ijt}}{q_{cjt}}\right)^{\gamma-1} \left(\sum_{k=1}^{N_{jt}} \frac{l_{kjt}}{l_{ijt}}\right)^{\frac{\gamma-\eta}{\eta-1}}, \forall i \in \{1,...,N_{jt}\}, \forall j \in [0,1] \tag{E.11}$$

The first equation is the relative labor allocation between two superstars i and k. The second equation is the allocation between superstar i and the fringe. Combined with (E.9), some algebra shows:

$$l_{ijt} = \frac{\hat{\omega}_{ijt}^{\eta} \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt}^{\eta} \left(\frac{q_{kjt}}{q_{ijt}}\right)^{\eta-1}\right)^{\frac{\gamma-\eta}{\eta-1}}}{\left(\frac{q_{cjt}}{q_{ijt}}\right)^{\gamma-1} + \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt}^{\eta} \left(\frac{q_{kjt}}{q_{ijt}}\right)^{\eta-1}\right)^{\frac{\gamma-1}{\eta-1}}}, \quad \forall i \in \{1,...,N_{jt}\}, \forall j \in [0,1] \tag{E.12}$$

$$l_{cjt} = \frac{1}{1 + \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt}^{\eta} \left(\frac{q_{kjt}}{q_{cjt}}\right)^{\eta-1}\right)^{\frac{\gamma-1}{\eta-1}}}, \quad \forall j \in [0,1] \tag{E.13}$$

Under the socially optimal choice of labor, aggregate log-output is:

$$\ln(Y_t) = \int_0^1 \ln(q_{cjt})\, dj + \frac{1}{\gamma - 1} \int_0^1 \ln\left(1 + \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt}^{\eta} \left(\frac{q_{ijt}}{q_{cjt}}\right)^{\eta-1}\right)^{\frac{\gamma-1}{\eta-1}}\right) dj \tag{E.14}$$

Next, we characterize optimal advertising. The first order condition for ω_{ijt} is:

$$\left(\frac{\eta}{\eta-1}\right) \frac{\left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} y_{kjt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{y_{cjt}^{\frac{\gamma-1}{\gamma}} + \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} y_{kjt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}} \frac{N_{jt}}{\left(\sum_{k=1}^{N_{jt}} (1+\omega_{kjt})\right)^2} \times \dots$$

$$\dots \times \left[y_{ijt}^{\frac{\eta-1}{\eta}} \sum_{h \neq i} (1+\omega_{hjt}) - \sum_{h \neq i} (1+\omega_{hjt}) y_{hjt}^{\frac{\eta-1}{\eta}} \right] = \frac{\chi_a \phi_a \omega_{ijt}^{\phi_a - 1}}{1 - A_t} \tag{E.15}$$

where A_t is the advertising share of GDP, defined in equation (E.2c). This can be written in terms of the labor choices of the planner (which were derived above):

$$\left(\frac{\eta}{\eta-1}\right) \frac{l_{ijt}}{\sum_{k=1}^{N_{jt}} (1+\omega_{ijt})} \left[\frac{\sum_{h \neq i} (1+\omega_{hjt})}{1+\omega_{ijt}} - \sum_{h \neq i} \frac{l_{hjt}}{l_{ijt}} \right] = \frac{\chi_a \phi_a \omega_{ijt}^{\phi_a - 1}}{1 - A_t} \tag{E.16}$$

E.3 Comparing the Planner's and the Decentralized Static Solutions

We now compare the planner's allocation of labor and advertising expenditures to the one from the decentralized economy (DE). We start with the labor choice. Labor demands can be written as:

$$l_{ijt}^{DE} = \sigma_{ijt} \left(\frac{M_{ijt}}{M_t}\right)^{-1} \quad \text{and} \quad l_{cjt}^{DE} = \sigma_{cjt} \left(\frac{M_{cjt}}{M_t}\right)^{-1} \tag{E.17}$$

In equation (E.17), M_t is the aggregate markup defined as a harmonic sales-weighted mean of firm-level markups:

$$M_t \equiv \left[\int_0^1 \left(\sigma_{cjt} + \sum_{i=1}^{N_{jt}} \sigma_{ijt} M_{ijt}^{-1} \right) dj \right]^{-1}$$

Further, one can show that the market shares (defined in equation (13)) can be written in terms of markups as follows:

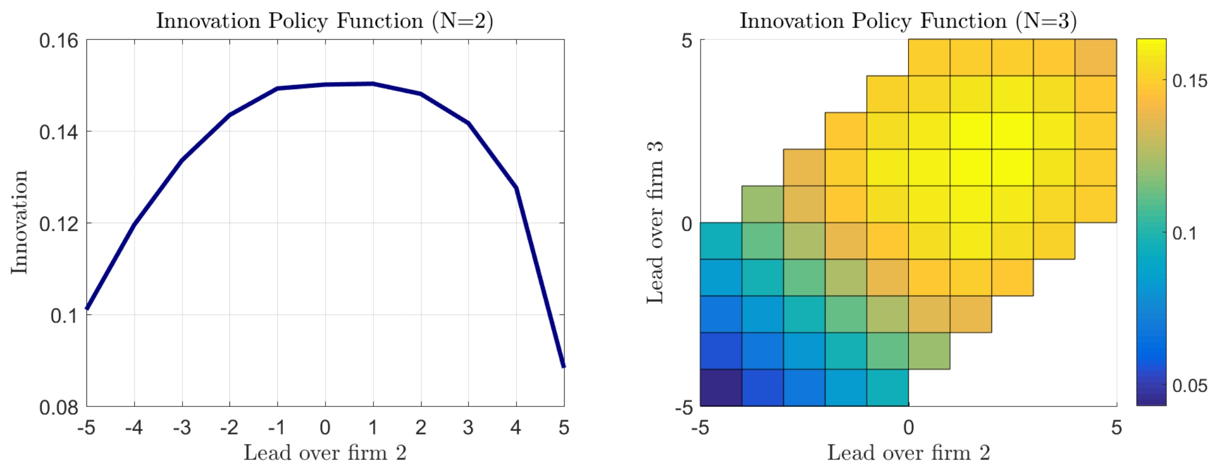
$$\sigma_{ijt} = \frac{\hat{\omega}_{ijt}^{\eta} \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt}^{\eta} \left(\frac{q_{kjt}}{q_{ijt}} \right)^{\eta-1} \left(\frac{M_{ijt}}{M_{kjt}} \right)^{\eta-1} \right)^{\frac{\gamma-\eta}{\eta-1}}}{\left(\frac{q_{cjt}}{q_{ijt}} \right)^{\gamma-1} \left(\frac{M_{ijt}}{M_{cjt}} \right)^{\gamma-1} + \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt}^{\eta} \left(\frac{q_{kjt}}{q_{ijt}} \right)^{\eta-1} \left(\frac{M_{ijt}}{M_{kjt}} \right)^{\eta-1} \right)^{\frac{\gamma-1}{\eta-1}}} \tag{E.18}$$

$$\sigma_{cjt} = \frac{1}{1 + \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt}^{\eta} \left(\frac{q_{kjt}}{q_{cjt}} \right)^{\eta-1} \left(\frac{M_{cjt}}{M_{kjt}} \right)^{\eta-1} \right)^{\frac{\gamma-1}{\eta-1}}} \tag{E.19}$$

Comparing allocation (E.17) with the social planner's (equations (E.12)-(E.13)), the only differences are the terms involving ratios of markups. Therefore, the two allocations coincide when $M_{ijt} = M_{kjt} = M_{cjt}$, $\forall i, k, j$. As $M_{cjt} = 1, \forall j$, by assumption, this means $M_t = 1$. In words, the labor allocation in the DE coincides with the planner's when all firms set zero markups. Otherwise, there is both within- and across-industry misallocation (indeed, recall that the planner allocates equal labor to all industries).

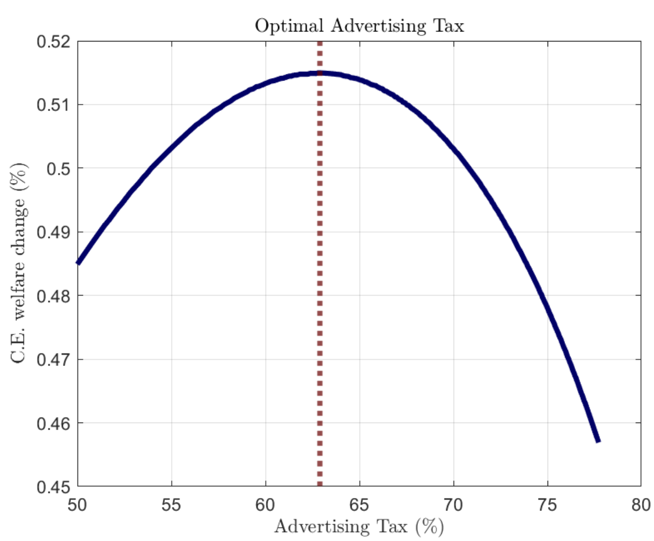
F Additional Figures and Tables

FIGURE F.1: INNOVATION POLICY FUNCTION



Notes: This figure presents the policy functions for innovation for the case of industries with $N = 2$ superstar firms (left panel) and $N = 3$ superstar firms (right panel). These policy functions are plotted from the perspective of a given firm, as functions of this firm's technological lead relative to its competitor(s), where a negative number means that the firm is lagging relative to its competitor. Firms innovate the most when they are close to being neck-to-neck, and innovation incentives decrease the higher the technological gap with their competitors.

FIGURE F.2: THE DYNAMIC WELFARE IMPACT OF ADVERTISING TAXES



Notes: This figure depicts the dynamic welfare impact of taxing advertising. The dotted line indicates the level of advertising tax which maximizes the change in consumption-equivalent welfare.

Table F.1: Markups, Advertising, and Innovation at the Firm Level

	Markup	R&D	Advertising	SG&A	Profitability
Markup	1.000				
R&D	0.359	1.000			
Advertising	0.698	0.853	1.000		
SG&A Expense	0.555	0.960	0.965	1.000	
Profitability	0.603	0.617	0.643	0.655	1.000

Notes: This table reports the correlation between markups, R&D expenditures, advertising expenditures, SG&A expenses, and profitability at the firm level.

Table F.2: CR4 Distributions: Data vs. Model

CR4	Data	Model
Mean	48.41%	43.13%
25th percentile	35.20%	38.43%
50th percentile	46.86%	43.26%
75th percentile	54.22%	48.79%

Notes: This table reports the distribution of four-firm concentration ratio (CR4), representing the market share of the four largest firms in each industry, both in the data and the model. The CR4 in the data is calculated based on all 3-digit BEA industries between 1976-2004 using Compustat data for top firms. The CR4 in the model is calculated based on the parameter estimates of the data sample (1976-2004). All statistics are calculated using total industry sales as weights.

Table F.3: Advertising and Relative Sales
(controlling for customer base accumulation)

	log advertising expenses
relative sales	5.891
	(0.181)***
relative sales sq.	-5.614
	(0.231)***
R^2	0.75
N	37,491

Notes: Robust asymptotic standard errors reported in parentheses are clustered at the firm level. The sample period is from 1976 to 2004 at the annual frequency. The regression controls for profitability, leverage, market-to-book ratio, log R&D stock, firm age, a persistent demand stock variable constructed following Fitzgerald *et al.* (2023) (using equation (2) in Fitzgerald *et al.* (2023) with their parameter values), the coefficient of variation of the firm's stock price, the number of firms in the industry, and a full set of year and SIC4 industry fixed effects. $^{***} = p < 0.01, ^{**} = p < 0.05, ^{*} = p < 0.1$.

Table F.4: The Dynamic Impact of Advertising Taxes and Subsidies on Macroeconomic Aggregates

	Benchmark	25% Tax	% change	50% Tax	% change	75% Tax	% change
Growth rate	2.201%	2.205%	0.17%	2.211%	0.44%	2.221%	0.92%
R&D/GDP	2.467%	2.463%	-0.18%	2.463%	-0.20%	2.473%	0.22%
Advertising/GDP (after-tax)	2.208%	2.080%	-5.80%	1.903%	-13.79%	1.615%	-26.85%
Average markup	1.342	1.335	-0.52%	1.326	-1.20%	1.312	-2.22%
Std. dev. markup	0.442	0.435	-1.71%	0.425	-3.97%	0.409	-7.42%
Labor share	0.638	0.640	0.30%	0.643	0.68%	0.646	1.27%
Average profitability	0.136	0.135	-0.80%	0.134	-1.80%	0.132	-3.22%
Average leader relative quality	0.510	0.508	-0.44%	0.504	-1.17%	0.497	-2.59%
Std. dev. leader relative quality	0.164	0.164	-0.19%	0.163	-0.64%	0.161	-1.74%
Superstar innovation	0.339	0.341	0.55%	0.344	1.46%	0.350	3.30%
Small firm innovation	0.096	0.097	0.80%	0.098	2.06%	0.101	4.46%
Output share of superstars	0.431	0.430	-0.42%	0.427	-0.89%	0.425	-1.46%
Average superstars per industry	2.864	2.877	0.45%	2.899	1.22%	2.944	2.80%
Mass of small firms	1.000	1.011	1.05%	1.028	2.75%	1.061	6.14%
Initial output	1.159	1.153	-0.47%	1.146	-1.06%	1.137	-1.87%
C.E. welfare change		0.300%		0.485%		0.478%	

	Optimal Tax (62.9%)	% change	20% Subsidy	% change	30% Subsidy	% change
Growth rate	2.215%	0.64%	2.198%	-0.12%	2.197%	-0.19%
R&D/GDP	2.466%	-0.08%	2.474%	0.25%	2.479%	0.46%
Advertising/GDP (after-tax)	1.777%	-19.52%	2.308%	4.56%	2.369%	7.31%
Average markup	1.320	-1.66%	1.348	0.43%	1.351	0.70%
Std. dev. markup	0.418	-5.51%	0.448	1.40%	0.452	2.28%
Labor share	0.644	0.95%	0.637	-0.24%	0.636	-0.40%
Average profitability	0.133	-2.46%	0.137	0.68%	0.138	1.11%
Average leader relative quality	0.501	-1.76%	0.511	0.29%	0.512	0.44%
Std. dev. leader relative quality	0.162	-1.07%	0.164	0.06%	0.164	0.04%
Superstar innovation	0.346	2.22%	0.338	-0.36%	0.337	-0.54%
Small firm innovation	0.099	3.07%	0.096	-0.56%	0.096	-0.86%
Output share of superstars	0.426	-1.17%	0.433	0.37%	0.434	0.61%
Average superstars per industry	2.917	1.87%	2.855	-0.29%	2.851	-0.43%
Mass of small firms	1.042	4.15%	0.993	-0.72%	0.989	-1.11%
Initial output	1.142	-1.44%	1.163	0.41%	1.166	0.67%
C.E. welfare change	0.515%		-0.381%		-0.691%	

Notes: This table reports the results of our policy experiment for different values of taxes and subsidies. The revenues from taxes are rebated back to the consumers, and subsidies are financed through lump-sum taxes.

TABLE F.5: OPTIMAL ADVERTISING TAX WITH DECEPTIVE ADVERTISING ($\delta = 1$)

	Benchmark	Optimal Tax (89.3%)	% change
Growth rate	2.201%	2.234%	1.50%
R&D/GDP	2.467%	2.495%	1.14%
Advertising/GDP (after-tax)	2.208%	1.291%	-41.51%
Average markup	1.342	1.298	-3.25%
Std. dev. markup	0.442	0.394	-11.02%
Labor share	0.638	0.650	1.86%
Average profitability	0.136	0.130	-4.54%
Average leader relative quality	0.510	0.487	-4.43%
Std. dev. leader relative quality	0.164	0.158	-3.37%
Superstar innovation	0.339	0.358	5.74%
Small firm innovation	0.096	0.104	7.40%
Output share of superstars	0.431	0.423	-1.85%
Average superstars per industry	2.864	3.004	4.91%
Mass of small firms	1.000	1.107	10.71%
Initial output	1.159	1.128	-2.62%
C.E. welfare change		1.201%	

Notes: This table presents the changes in the relevant economic variables under the optimal advertising tax rate compared to the baseline economy in the extended model with fully deceptive advertising ($\delta = 1$).

TABLE F.6: EXTENDED MODEL PARAMETERS AND TARGET MOMENTS
(NON-COMBATIVE ADVERTISING WITH $\Lambda = 1$)

A. Calibrated parameters

Parameter	*Description*	*Value*
λ	Innovation step size	0.1830
η	Elasticity within industry	13.1154
γ	Elasticity between superstars and fringe	1.6696
χ	Superstar cost scale	75.6619
ν	Small firm cost scale	2.4681
ζ	Competitive fringe ratio	1.2765
ϕ	Superstar cost convexity	4.2645
ϵ	Small firm cost convexity	4.3789
τ	Small firm exit rate	0.1151
ψ	Entry cost scale	0.0728
χ_a	Advertising cost scale	0.0929
ϕ_a	Advertising cost convexity	3.9878

B. Moments

Target moments	*Data*	*Model*
Growth rate	2.204%	2.206%
R&D/GDP	2.435%	2.289%
Advertising/GDP	2.200%	2.213%
Average markup	1.350	1.370
Standard deviation of markups	0.346	0.575
Labor share	0.652	0.645
Firm entry rate	0.115	0.115
Average profitability	0.144	0.130
Average leader relative quality	0.749	0.521
Standard deviation of leader relative quality	0.223	0.165
Regression (1a), linear coefficient: β_1^{inn}	0.629	1.043
Regression (1a), top point: $-\beta_1^{inn}/(2\beta_2^{inn})$	0.505	0.475
Regression (1b), linear coefficient: β_1^{adv}	6.260	7.544
Regression (1b), top point: $-\beta_1^{adv}/(2\beta_2^{adv})$	0.533	0.563

Notes: Panel A reports the calibrated parameters. Panel B reports the simulated and empirical moments.

Table F.7: The Impact of Advertising Shutdown and Optimal Advertising Tax (Non-combative Advertising with $\Lambda = 1$)

	Benchmark	Shutdown	% change	Optimal Tax (28.6%)	% change
Growth rate	2.206%	2.248%	1.91%	2.209%	0.16%
R&D/GDP	2.289%	2.411%	5.36%	2.296%	0.34%
Advertising/GDP (after-tax)	2.213%	0.000	-100.00%	2.054%	-7.16%
Average markup	1.370	1.286	-6.09%	1.364	-0.43%
Std. dev. markup	0.575	0.474	-17.56%	0.569	-1.13%
Labor share	0.645	0.666	3.29%	0.647	0.23%
Average profitability	0.130	0.124	-4.53%	0.129	-0.31%
Average leader relative quality	0.521	0.473	-9.22%	0.518	-0.55%
Std. dev. leader relative quality	0.165	0.150	-9.12%	0.164	-0.54%
Superstar innovation	0.304	0.341	12.22%	0.306	0.68%
Small firm innovation	0.095	0.107	12.53%	0.096	0.76%
Output share of superstars	0.378	0.355	-6.02%	0.376	-0.54%
Average superstars per industry	2.867	3.194	11.44%	2.884	0.62%
Mass of small firms	1.000	1.217	21.73%	1.011	1.07%
Initial output	3.148	2.886	-8.33%	3.125	-0.73%
C.E. welfare change		-5.381%		0.119%	

Notes: This table presents the changes in the relevant economic variables under the advertising shutdown and optimal advertising tax experiments compared to the baseline economy in the extended model with fully non-combative advertising ($\Lambda = 1$).

TABLE F.8: EXTENDED MODEL PARAMETERS AND TARGET MOMENTS
(MODEL WITH INTERNALLY CALIBRATED Λ)

A. Calibrated parameters

Parameter	*Description*	*Value*
λ	Innovation step size	0.1706
η	Elasticity within industry	11.9053
γ	Elasticity between superstars and fringe	3.0215
χ	Superstar cost scale	77.1575
ν	Small firm cost scale	3.1016
ζ	Competitive fringe ratio	0.7218
ϕ	Superstar cost convexity	4.4855
ϵ	Small firm cost convexity	4.5392
τ	Small firm exit rate	0.1151
ψ	Entry cost scale	0.0528
χ_a	Advertising cost scale	0.0583
ϕ_a	Advertising cost convexity	3.3062
Λ	Degree of advertising non-combativeness	0.0199

B. Moments

Target moments	*Data*	*Model*
Growth rate	2.204%	2.236%
R&D/GDP	2.435%	2.417%
Advertising/GDP	2.200%	2.211%
Average markup	1.350	1.344
Standard deviation of markups	0.346	0.451
Labor share	0.652	0.639
Firm entry rate	0.115	0.115
Average profitability	0.144	0.136
Average leader relative quality	0.749	0.516
Standard deviation of leader relative quality	0.223	0.162
Regression (1a), linear coefficient: β_1^{inn}	0.629	0.997
Regression (1a), top point: $-\beta_1^{inn}/(2\beta_2^{inn})$	0.505	0.488
Regression (1b), linear coefficient: β_1^{adv}	6.260	7.186
Regression (1b), top point: $-\beta_1^{adv}/(2\beta_2^{adv})$	0.533	0.592
Regression (2): β_3/β_4	-1.010	-1.144

Notes: Panel A reports the calibrated parameters. Panel B reports the simulated and empirical moments.

Table F.9: The Impact of Advertising Shutdown and Optimal Advertising Tax (Model With Internally Calibrated $\Lambda = 0.0199$)

	Benchmark	Shutdown	% change	Optimal Tax (53.99%)	% change
Growth rate	2.236%	2.317%	3.62%	2.249%	0.57%
R&D/GDP	2.417%	2.517%	4.11%	2.395%	-0.91%
Advertising/GDP (after-tax)	2.211%	0.000	-100.00%	1.867%	-15.57%
Average markup	1.344	1.247	-7.19%	1.323	-1.53%
Std. dev. markup	0.451	0.339	-24.94%	0.429	-5.01%
Labor share	0.639	0.666	4.21%	0.644	0.87%
Average profitability	0.136	0.123	-9.49%	0.133	-2.50%
Average leader relative quality	0.516	0.457	-11.45%	0.511	-1.03%
Std. dev. leader relative quality	0.162	0.146	-10.14%	0.163	0.21%
Superstar innovation	0.333	0.386	15.98%	0.338	1.44%
Small firm innovation	0.095	0.111	17.37%	0.097	2.41%
Output share of superstars	0.423	0.409	-3.37%	0.418	-1.31%
Average superstars per industry	2.847	3.232	13.52%	2.878	1.08%
Mass of small firms	1.000	1.325	32.47%	1.030	3.02%
Initial output	1.163	1.101	-5.34%	1.147	-1.37%
C.E. welfare change	-	-1.316%	-	0.364%	-

Notes: This table presents the changes in the relevant economic variables under the advertising shutdown and optimal advertising tax experiments compared to the baseline economy in the extended model with calibrated advertising non-combativeness parameter, Λ.

TABLE F.10: EXTENDED MODEL PARAMETERS AND TARGET MOMENTS (BERTRAND COMPETITION MODEL)

A. Calibrated parameters

Parameter	*Description*	*Value*
λ	Innovation step size	0.2492
η	Elasticity within industry	3.2408
γ	Elasticity between superstars and fringe	3.2508
χ	Superstar cost scale	62.451
ν	Small firm cost scale	3.5236
ζ	Competitive fringe ratio	0.8126
ϕ	Superstar cost convexity	3.7648
ϵ	Small firm cost convexity	3.6111
τ	Small firm exit rate	0.1151
ψ	Entry cost scale	0.1238
χ_a	Advertising cost scale	0.2873
ϕ_a	Advertising cost convexity	5.0222

B. Moments

Target moments	*Data*	*Model*
Growth rate	2.204%	2.229%
R&D/GDP	2.435%	2.364%
Advertising/GDP	2.200%	2.301%
Average markup	1.350	1.306
Standard deviation of markups	0.346	0.311
Labor share	0.652	0.633
Firm entry rate	0.115	0.115
Average profitability	0.144	0.144
Average leader relative quality	0.749	0.489
Standard deviation of leader relative quality	0.223	0.136
Regression (1a), linear coefficient: β_1^{inn}	0.629	0.821
Regression (1a), top point: $-\beta_1^{inn}/(2\beta_2^{inn})$	0.505	0.433
Regression (1b), linear coefficient: β_1^{adv}	6.260	8.581
Regression (1b), top point: $-\beta_1^{adv}/(2\beta_2^{adv})$	0.533	0.499

Notes: Panel A reports the calibrated parameters. Panel B reports the simulated and empirical moments.

Table F.11: The Impact of Advertising Shutdown and Optimal Advertising Tax (Bertrand Competition Model)

	Benchmark	Shutdown	% change	Optimal Tax (90.65%)	% change
Growth rate	2.229%	2.238%	0.42%	2.241%	0.54%
R&D/GDP	2.364%	2.606%	10.24%	2.453%	3.77%
Advertising/GDP (after-tax)	2.301%	0.000	-100.00%	1.588%	-31.01%
Average markup	1.306	1.283	-1.73%	1.297	-0.67%
Std. dev. markup	0.311	0.289	-7.14%	0.303	-2.70%
Labor share	0.633	0.641	1.14%	0.636	0.43%
Average profitability	0.144	0.156	8.39%	0.147	2.06%
Average leader relative quality	0.489	0.444	-9.18%	0.473	-3.34%
Std. dev. leader relative quality	0.136	0.117	-13.48%	0.130	-4.00%
Superstar innovation	0.276	0.306	10.83%	0.287	4.10%
Small firm innovation	0.069	0.064	-6.70%	0.069	-0.20%
Output share of superstars	0.499	0.495	-0.82%	0.498	-0.34%
Average superstars per industry	3.291	3.563	8.28%	3.394	3.13%
Mass of small firms	1.000	1.187	18.72%	1.072	7.21%
Initial output	1.368	1.348	-1.47%	1.360	-0.59%
C.E. welfare change		0.830%		1.853%	

Notes: This table presents the changes in the relevant economic variables under the advertising shutdown and optimal advertising tax experiments compared to the baseline economy in the extended model with Bertrand competition.

TABLE F.12: BENCHMARK MODEL PARAMETERS AND TARGET MOMENTS
(CALIBRATION WITH $\bar{n} = 6$)

A. Calibrated parameters

Parameter	*Description*	*Value*
λ	Innovation step size	0.1401
η	Elasticity within industry	10.8186
γ	Elasticity between superstars and fringe	3.0849
χ	Superstar cost scale	78.1071
ν	Small firm cost scale	3.1077
ζ	Competitive fringe ratio	0.6821
ϕ	Superstar cost convexity	4.9156
ϵ	Small firm cost convexity	5.0898
τ	Small firm exit rate	0.1151
ψ	Entry cost scale	0.0548
χ_a	Advertising cost scale	0.0672
ϕ_a	Advertising cost convexity	3.6759

B. Moments

Target moments	*Data*	*Model*
Growth rate	2.204%	2.207%
R&D/GDP	2.435%	2.363%
Advertising/GDP	2.200%	2.205%
Average markup	1.350	1.350
Standard deviation of markups	0.346	0.437
Labor share	0.652	0.634
Firm entry rate	0.115	0.115
Average profitability	0.144	0.143
Average leader relative quality	0.749	0.498
Standard deviation of leader relative quality	0.223	0.154
Regression (1a), linear coefficient: β_1^{inn}	0.629	1.043
Regression (1a), top point: $-\beta_1^{inn}/(2\beta_2^{inn})$	0.505	0.483
Regression (1b), linear coefficient: β_1^{adv}	6.260	7.750
Regression (1b), top point: $-\beta_1^{adv}/(2\beta_2^{adv})$	0.533	0.519

Notes: The baseline model is re-calibrated with the maximum number of productivity steps between any two superstar firms, $\bar{n}$, set to 6. Panel A reports the calibrated parameters. Panel B reports the simulated and actual moments.

Table F.13: Advertising Shutdown: The Dynamic Impact on Macroeconomic Aggregates (Calibration with $\bar{n} = 6$)

	Benchmark	Shutdown	% change
Growth rate	2.207%	2.268%	2.77%
R&D/GDP	2.363%	2.520%	6.61%
Advertising/GDP	2.205%	0.000	-100.00%
Average markup	1.350	1.262	-6.47%
Std. dev. markup	0.437	0.340	-22.22%
Labor share	0.634	0.658	3.87%
Average profitability	0.143	0.133	-7.36%
Average leader relative quality	0.498	0.439	-11.94%
Std. dev. leader relative quality	0.154	0.134	-12.55%
Superstar innovation	0.410	0.478	16.65%
Small firm innovation	0.107	0.123	14.99%
Output share of superstars	0.448	0.439	-2.16%
Average superstars per industry	2.935	3.352	14.22%
Mass of small firms	1.000	1.361	36.06%
Initial output	1.123	1.070	-4.67%
C.E. welfare change	-	-1.187%	-

Notes: This table presents the changes in the relevant macroeconomic aggregates under the advertising shutdown compared to the baseline economy with the maximum number of productivity steps between any two superstar firms, $\bar{n}$, set to 6.

TABLE F.14: BENCHMARK MODEL PARAMETERS AND TARGET MOMENTS (CALIBRATION WITH $\bar{N} = 5$)

A. Calibrated parameters

Parameter	*Description*	*Value*
λ	Innovation step size	0.1643
η	Elasticity within industry	11.6863
γ	Elasticity between superstars and fringe	2.9636
χ	Superstar cost scale	77.3020
ν	Small firm cost scale	3.1650
ζ	Competitive fringe ratio	0.7053
ϕ	Superstar cost convexity	4.4879
ϵ	Small firm cost convexity	4.5288
τ	Small firm exit rate	0.1151
ψ	Entry cost scale	0.0642
χ_a	Advertising cost scale	0.0730
ϕ_a	Advertising cost convexity	3.3544

B. Moments

Target moments	*Data*	*Model*
Growth rate	2.204%	2.211%
R&D/GDP	2.435%	2.558%
Advertising/GDP	2.200%	2.264%
Average markup	1.350	1.342
Standard deviation of markups	0.346	0.437
Labor share	0.652	0.637
Firm entry rate	0.115	0.115
Average profitability	0.144	0.136
Average leader relative quality	0.749	0.496
Standard deviation of leader relative quality	0.223	0.172
Regression (1a), linear coefficient: β_1^{inn}	0.629	1.005
Regression (1a), top point: $-\beta_1^{inn}/(2\beta_2^{inn})$	0.505	0.480
Regression (1b), linear coefficient: β_1^{adv}	6.260	7.753
Regression (1b), top point: $-\beta_1^{adv}/(2\beta_2^{adv})$	0.533	0.512

Notes: The baseline model is re-calibrated with the maximum number of superstar firms in an industry $\bar{N}$ set to 5. Panel A reports the calibrated parameters. Panel B reports the simulated and actual moments.

TABLE F.15: THE IMPACT OF ADVERTISING SHUTDOWN AND OPTIMAL ADVERTISING TAX (CALIBRATION WITH $\bar{N} = 5$)

	Benchmark	Shutdown	% change	Optimal Tax (57.55%)	% change
Growth rate	2.211%	2.288%	3.48%	2.221%	0.47%
R&D/GDP	2.558%	2.743%	7.24%	2.554%	-0.14%
Advertising/GDP (after-tax)	2.264%	0.000	-100.00%	1.868%	-17.52%
Average markup	1.342	1.251	-6.77%	1.322	-1.48%
Std. dev. markup	0.437	0.335	-23.23%	0.416	-4.74%
Labor share	0.637	0.663	4.06%	0.643	0.87%
Average profitability	0.136	0.124	-8.72%	0.133	-2.23%
Average leader relative quality	0.496	0.428	-13.69%	0.488	-1.56%
Std. dev. leader relative quality	0.172	0.157	-8.46%	0.171	-0.46%
Superstar innovation	0.356	0.425	19.48%	0.363	1.97%
Small firm innovation	0.114	0.148	29.76%	0.118	3.50%
Output share of superstars	0.438	0.428	-2.33%	0.433	-1.16%
Average superstars per industry	3.035	3.604	18.77%	3.091	1.86%
Mass of small firms	1.000	1.330	33.03%	1.032	3.22%
Initial output	1.164	1.106	-4.96%	1.148	-1.38%
C.E. welfare change	-	-1.061%	-	0.405%	-

Notes: This table presents the changes in the relevant economic variables under the advertising shutdown and optimal advertising tax experiments compared to the baseline economy in the recalibrated baseline model with the maximum number of superstar firms in an industry $\bar{N}$ set to 5.