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

Laurent Cavenaile Murat Alp Celik Pau Roldan-Blanco Xu Tian

August 6, 2023

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.

Keywords: innovation, advertising, markups, growth, industry dynamics, misallocation, business dynamism.

JEL Classification: E20, G30, M30, O30, O40.

*Contact Information: Laurent Cavenaile, Rotman School of Management, University of Toronto, Toronto, ON M5S 3E6. Email: laurent.cavenaile@utoronto.ca; Murat Alp Celik, 150 St. George Street, University of Toronto, Toronto, ON M5S 3G7. Email: murat.celik@utoronto.ca; Pau Roldan-Blanco, Banco de España and CEMFI, Calle Alcalá 48, 28014, Madrid. Email: pau.roldan@bde.es; Xu Tian, Department of Finance, Terry College of Business, University of Georgia, 620 South Lumpkin Street, Athens, GA 30602. Email: xu.tian@uga.edu. We would like to thank David Argente, Salome Baslandze, Maarten De Ridder, Javier Fernandez, Giammario Impullitti, Charles Jones, Michael Klein, Omar Licandro, Yueyuan Ma, Atif Mian, Sara Moreira, Pedro Monteiro, Sergio Ocampo, Miquel Oliver i Vert, Jesse Perla, Thomas Philippon, Michael Peters, Nicolas Sahuguet, Kenneth Simons, Christopher Tonetti, Nicolas Vincent, Toni Whited, Peifan Wu, Daniel Yi Xu, Mehmet Yorukoglu, and conference and seminar participants at NFA Annual Meeting 2021, Banco de España 4th Annual Research Conference, University of Nottingham, FMA Annual Meeting 2021, CMSG Annual Meeting 2021, Laval University, Universitat Autònoma de Barcelona, Rensselaer Polytechnic Institute, Bristol University, EWMES 2021, HEC Montreal, BSE Summer Forum 2022, SED Meetings 2022, CEA Annual Meetings 2022, EARIE Conference 2022, University of Toronto, University of Liège, AFA Annual Meeting 2023, UQAM, and Bank of Canada for their helpful comments. The views expressed in this paper are those of the authors and do not necessarily coincide with those of the Banco de España or the Eurosystem. Celik gratefully acknowledges financial support from SSHRC Insight Development Grant (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.

1 Introduction

Firms that compete against their peers have several ways to improve their profits, markups, and market shares. Innovation – spending resources on research and development (R&D) to come up with new products or more efficient technologies – is a well-studied one, which is also considered to be the engine of growth in developed economies. Advertising is another activity through which firms can achieve the same desired outcomes, albeit without directly contributing to long-run productivity growth. Indeed, firms might spend exorbitant amounts on advertising in response to the advertising efforts of their competitors, leading to an inefficient "rat race" equilibrium with excessive spending in the aggregate. Since both activities serve similar purposes, firms' decisions to innovate and to advertise inexorably interact, within the firm itself, as well as across all the firms in the same industry.

In practice, firms devote significant amounts of resources to both innovation and advertising. Since 1980, R&D accounts for 2.44% of GDP in the U.S., whereas advertising alone represents 2.20%.[1] In other words, as a society, we spend as much on developing new products and technologies ("substance") as on simply marketing them ("style"). For instance, Procter & Gamble Company spent 10.8% of its revenue on advertising between 2007 and 2016, which is quadruple the amount it spent on R&D (2.6%). For Unilever, the numbers were 13.3% and 2.0%, respectively. Two natural questions to ask are (1) whether the heavy spending on advertising is socially efficient, and (2) whether firms would focus more on innovation rather than advertising if the latter became more costly. Answering these questions and deriving their policy implications require a unified framework.

In this paper, we present a new model of firm and industry dynamics which can, in a single framework, elucidate the role of innovation and advertising for market concentration, markups, and productivity growth, offering a realistic representation of how these two forms of intangible investment interact and relate to competition at the aggregate level. In the model, the market structure of the economy is endogenous: the within-industry composition between small and large firms, as well as the number of large firms within each industry, is an equilibrium outcome. Market

[1] 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% according to Arkolakis (2010)).

structure is shaped, in turn, by the production, innovation, and advertising decisions of large firms, which determine their market share within the industry and the markups they charge to final consumers. This is because large firms behave strategically, internalizing the effects of their production, advertising, and innovation decisions on the industry's aggregate expenditures. Small firms, by contrast, are atomistic, 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.

This rich yet tractable setting allows us to tackle relevant policy questions. First, our quantitative model allows us to study the role of the interaction between innovation and advertising for static allocative efficiency, and to conduct policy counterfactuals to understand the role of intangibles for static (physical-input) misallocation. Second, we can also study the dynamic consequences of the innovation-advertising interaction for economic growth and industry dynamics. Static and dynamic considerations, as well as within- and between-industry dynamics, all matter for social welfare. Therefore, the model offers a broad set of endogenously-generated responses, allowing us to analyze optimal policy from an all-encompassing standpoint, in general equilibrium.

In the model, 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 on their own market share. Because large firms behave strategically, they must also internalize the effect of their production and advertising decisions on industry-level expenditures. 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

[2] In a model extension in Section 5, we relax this assumption and allow for advertising to be non-combative.

aggregate level has an impact on the rate of economic growth and social welfare.

To study these questions quantitatively, we estimate the model by the simulated method of moments to fit key empirical patterns relating advertising and innovation to competition. As our quantitative analysis focuses on the aggregate welfare implications stemming from both micro- and macro-level effects, in the estimation stage, we make sure that the model fits the data well at different levels of aggregation, namely between firms within industries, across industries, and in terms of macroeconomic aggregates. Importantly, our model can reproduce the empirically-observed non-linear relationship between innovation, advertising, and market share within industries, which helps us discipline our counterfactual experiments. In the data, both innovation and advertising expenditures exhibit an inverted-U shaped relationship with respect to a firm's relative sales within their industry. The estimated model matches the linear term and top point of both of these hump-shaped curves. Using the estimated set of parameters, we find that markups, R&D expenditures, and advertising expenditures are all positively correlated at the firm level, consistent with the idea that firms use both types of intangibles to increase their profits and harness greater market power.

Next, we conduct a series of counterfactual experiments to understand the interplay between R&D and advertising at various levels of aggregation, and ultimately to assess the effects of advertising on misallocation, growth, and welfare. In the first experiment, we compare the estimated model with a counterfactual economy in which advertising is shut down completely (e.g., it is infinitely costly for firms). We find that shutting down advertising increases firm-level investment in R&D, both by large firms as well as small firms, thereby increasing both aggregate innovation and the rate of economic growth. Thus, advertising and R&D are substitutes in our estimated economy, consistent with the empirical findings in Cavenaile and Roldan-Blanco (2021), providing an out-of-sample validation test for the model.[3]

Shutting down advertising also affects markups and allocative efficiency through changes in the competitive structure of industries. We find that the average net markup decreases by one quarter of its value relative to the baseline economy, as large firms cease to use advertising as a tool to shift demand and profits toward their products. This implies that advertising is responsible for

[3]Note that, under different parameter values, advertising and R&D can be complements rather than substitutes. Therefore, this is a quantitative result rather than a theoretical implication of the framework.

a significant fraction of the empirically-observed average markup. Because there is less product differentiation, markups are lower and the labor share higher in an economy without advertising. However, in spite of these effects, we find that advertising in fact improves allocative efficiency because of its reallocative effects. While increasing markups, advertising simultaneously reallocates physical inputs away from the less efficient firms, and towards the more efficient industry leaders. It simultaneously amplifies the relative perceived quality of the more abundant and cheaper-to-produce varieties. While markups themselves lower efficiency, the latter two effects quantitatively dominate.

To assess the relative quantitative importance for social welfare of the various static and dynamic channels identified above, we show that the change in welfare can be decomposed into changes in relative wages, the relative industry output of large firms, the consumption share of GDP, and the rate of economic growth.[4] We find substantial differences between static and dynamic welfare changes. Statically (i.e., without adjustments in the firm productivity 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 by allowing the distribution to adjust undoes some of these losses in the long run, as shutting down advertising also raises the consumption share of GDP and increases the rate of economic growth through the substitution effect between R&D and advertising. On the net, the combination of the various static and dynamic conflicting forces results in a welfare loss of 0.86% in consumption-equivalent terms from shutting down advertising.

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, and at the considerably high rate of 62.9%. How does one reconcile the two findings? The answer lies in understanding how taxation differs from a complete shutdown. 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

[4]While the consumption-equivalent 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. Furthermore, the extensions in Section 5 consider two alternatives for normative implications.

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, the taxes reduce the excessive spending on advertising due to the "rat race" between the superstars, while still largely preserving the relative market shares in equilibrium. This makes advertising an ideal candidate for taxation to raise revenues while simultaneously increasing dynamic efficiency.

We also investigate the heterogeneous impact of our counterfactual experiments on the value and market share of firms. We find that the most productive superstars (leaders) are adversely affected by both the advertising shutdown and taxes, whereas the less productive superstars (followers), as well as small firms and entrepreneurs, all benefit. This elucidates how both experiments positively affect business dynamism.

One might reasonably wonder whether our quantitative findings are contingent on the specific assumptions regarding how advertising enters consumer preferences, or the modeling choices regarding how the advertising efforts by firms affect demand shifters in the same industry. Motivated by such concerns, we build two extensions to our model, deceptive advertising and non-combative advertising, which can be interpreted as placing more weight on the persuasive and informative views of advertising compared to the baseline, respectively. We also consider an alternative model in which superstar firms compete in prices *à la* Bertrand instead of in quantities. Repeating the quantitative experiments under these extended models reveals that, even under the most extreme parametrizations, almost all of our main findings are preserved, such as the optimality of positive advertising taxes, the aggregate substitution between innovation and advertising, implications on business dynamism, the positive role of advertising in reducing static misallocation in oligopolistic markets, and its overall usefulness for raising welfare, demonstrating the robustness of our main conclusions.

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), Argente, Fitzgerald, Moreira, and Priolo (2021), Greenwood, Ma, and Yorukoglu (2022), Einav, Klenow, Levin, and Murciano-Goroff (2022), Ignaszak and Sedláček (2022), Pearce and Wu (2022), Dinlersoz, Goldschlag, Yorukoglu, and Zolas (2023), and Cavenaile, Celik, Perla, and Roldan-Blanco (2023)). The literature has investigated, for instance, how intangibles may be behind several trends related to business dynamism, market concentration and markups (e.g., De Ridder (2020), Weiss (2020), Cavenaile, Celik, and Tian (2021), Feijoo Moreira (2021), and Aghion, Bergeaud, Boppart, Klenow, and Li (2022)), or how they may affect markup cyclicality (e.g., Roldan-Blanco and Gilbukh (2021)), firm's market value and risk (e.g., Belo, Lin, and Vitorino (2014), Corhay, Kung, and Schmid (2020b)), asset pricing (e.g., Corhay, Kung, and Schmid (2020a), Dou, Ji, and Wu (2022)), 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, Haller, and Yedid-Levi (2022)). Argente, Fitzgerald, Moreira, and Priolo (2021) find that successful entrants in the consumer food sector build market share by reaching new customers in different geographical markets through product placement and direct advertising. They build and estimate a new structural model of endogenous customer base acquisition through marketing and advertising to match these facts.

Our model is most closely related to recent macroeconomic models with advertising such as Cavenaile, Celik, Perla, and Roldan-Blanco (2023), Baslandze, Greenwood, Marto, and Moreira (2023), Rachel (2022), Greenwood, Ma, and Yorukoglu (2022), Cavenaile and Roldan-Blanco (2021), and Klein and Şener (2023).[5] Cavenaile, Celik, Perla, and Roldan-Blanco (2023) build a model of targeted advertising in which consumers' awareness sets expand over the lifetime of an industry, resulting in a better consumer-product match. Baslandze, Greenwood, Marto, and Moreira (2023) focus on the advent of digital advertising, which is more targeted than traditional advertising, and on its growth and welfare implications through the consequent increase in product varieties. Rachel (2022) studies how the provision of free leisure-enhancing technologies which

[5] Ignaszak and Sedláček (2022) and Einav, Klenow, Levin, and Murciano-Goroff (2022) also consider how demand-side factors in the form of customer capital accumulation can affect firm dynamics and innovation in a model of endogenous growth with monopolistic competition. Afrouzi, Drenik, and Kim (2021) propose a model of customer accumulation through advertising in a model of monopolistic competition, and study the misallocation of customers across firms with exogenous variation in productivity.

firms use to build their brand equity (e.g., through advertising) can explain the observed decline in hours worked and negatively affect innovation and TFP growth. Greenwood, Ma, and Yorukoglu (2022) also consider the implication of advertising embedded in free media on hours worked and welfare, but do not study its effect on economic growth. They find that the expansion of free media arising from the advent of digital advertising is welfare improving. However, some advertising is wasteful and a tax on advertising might be required to correct for this source of inefficiency. In a model with monopolistic competition, Cavenaile and Roldan-Blanco (2021) show that there exists an interaction between R&D and advertising investment at the firm level which shapes the firm size distribution, firm dynamics, and long-run economic growth. Klein and Şener (2023) study how both informative and combative advertising affect the speed of diffusion of innovation. Investigating policy implications, they find that R&D subsidies increase innovation rates, but decrease advertising and diffusion, leading to an ambiguous effect on growth and welfare.

Our paper contributes to this body of work across several dimensions. The model we develop is at the intersection of four different strands of literature. It features oligopolistic competition as in Atkeson and Burstein (2008), but the productivities of firms are endogenously determined in a step-by-step innovation framework as in Aghion, Harris, Howitt, and Vickers (2001). Furthermore, there is endogenous entry and exit of large superstar firms, as well as an endogenous mass of small firms, as in Cavenaile, Celik, and Tian (2021), which results in an endogenous market structure found in dynamic industrial organization models such as Ericson and Pakes (1995), except in general equilibrium. In addition, we model advertising decisions at the firm level, which in equilibrium directly affect market shares and markups as well as the dynamic R&D decisions of all firms, and consequently the stationary distribution of different industry states and aggregate growth. This rich framework allows us to study how the interaction between R&D and advertising affects market concentration, markups, productivity growth, and welfare. In particular, we can study how advertising affects both static and dynamic efficiency, which leads to unveiling the role of advertising in affecting the misallocation of resources across oligopolistically competing firms. In our estimation for the United States, we find this effect to be large and positive. Dynamically, we can assess how advertising affects the incentives to innovate and long-run economic growth. As in Cavenaile and

Roldan-Blanco (2021), Rachel (2022) and Klein and Şener (2023), we find that advertising and R&D are substitutes, even though the theoretical frameworks are different. Compared to these papers, our model allows us to study in a single framework how advertising affects static resource misallocation between firms with different productivity, on top of its dynamic effect on firms' innovation decisions. As a result, the overall welfare effect of advertising is *a priori* ambiguous. We can also study the heterogeneous implications of advertising for leader vs. laggard firms.

More generally, in terms of methodology, our paper belongs to the growing literature that employs structural model calibration or estimation to address corporate finance questions in capital investment, leverage choice, governance, and valuation.[6] It 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)).[7] In the literature, advertising is usually modeled as a demand shifter. Following this tradition, we model advertising as a technology that shifts consumer preferences for certain goods to the detriment of competitors' products, which in equilibrium means that firms can use advertising expenditures to shift demand toward their own goods, thereby affecting their market share and the markups they set. To draw this connection between firm-level advertising and market structure, we rely on observations from various papers relating market concentration to intangible investments. Crouzet and Eberly (2019) argue that the increase in intangible capital investments driven by large firms may be behind the rise in industry concentration in the last two decades, while De Loecker, Eeckhout, and Unger (2020) find a positive firm-level relation between markups and both R&D and advertising expenditures for publicly-traded firms. In our model, these types of relationships emerge endogenously, and help explain the macroeconomic implications of R&D and advertising on growth and welfare.[8]

Outline The remainder of the paper is organized as follows: Section 2 presents our model of endogenous markups, innovation, advertising, and market structure. Section 3 discusses the estimation

[6] See Hennessy and Whited (2005), Taylor (2010), Jermann and Quadrini (2012), Nikolov and Whited (2014), Glover and Levine (2017), David (2021), David, Schmid, and Zeke (2022), Celik, Tian, and Wang (2022), and Terry (2023) among others.

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

[8] This feature of our model links our paper to the broader macroeconomics literature on competition and markups (see e.g., Covarrubias, Gutiérrez, and Philippon (2020), Gutiérrez, Jones, and Philippon (2021)).

of the model and the main quantitative features of the equilibrium. Section 4 conducts a number of counterfactual experiments using the estimated parameters, describing the macroeconomic effects of advertising on the composition of industries, the level and dispersion of markups, the rate of economic growth, and social welfare (Section 4.1). Moreover, we analyze the optimal taxation of advertising within the context of the estimated model (Section 4.2). Finally, we consider the heterogeneous impact of the counterfactual experiments on firm value and market share (Section 4.3). Section 5 presents the extended models with deceptive and non-combative advertising, and with Bertrand competition instead of Cournot. Section 6 concludes.

2 Model

2.1 Environment

In this section, we develop an oligopolistic general-equilibrium growth model with firm heterogeneity in which market structure is endogenous, and firms' production, innovation, and advertising decisions strategically interact. This new model of firm and industry dynamics can, in a single framework, study the role of innovation and advertising for market concentration, markups, and productivity growth, offering a realistic representation of how these two forms of intangible inputs interact and relate to competition at the aggregate level.

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

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

where $\rho > 0$ is the time discount rate, and C_t is consumption of the 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 of 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. The usual no-Ponzi-scheme condition holds.

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

$$Y_t = \exp\left(\int_0^1 \ln\left(y_{jt}\right)\ dj\right) \tag{2}$$

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, $N_{jt} \in \{1, ..., \bar{N}\}$, 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:

$$y_{jt} = \left(\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}} + \tilde{y}_{sjt}^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}} \tag{3}$$

where $\tilde{y}_{cjt}$ denotes the output of the fringe, $\tilde{y}_{sjt}$ denotes the output of superstars, and $\gamma \geq 1$ is the elasticity of substitution between the two. Fringe firms produce a homogeneous product, so:

$$\tilde{y}_{cjt} = \int_{F_{jt}} y_{ckjt}\ dk \tag{4}$$

where F_{jt} is the endogenous set of small firms in the fringe in industry j at time t. 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.[9] Total production by superstars of industry j at time t is given by:

$$\tilde{y}_{sjt} = \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt} y_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \tag{5}$$

[9] See Section 5 for the alternative model with Bertrand competition.

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

$$\hat{\omega}_{ijt} = \frac{1 + \omega_{ijt}}{\frac{1}{N_{jt}} \sum_{k=1}^{N_{jt}} (1 + \omega_{kjt})} \tag{6}$$

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 which allows firms to shift the perceived quality of their own product. 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 the same ω level, then all varieties would have the same baseline quality.[10] This baseline quality coincides with the quality of the product of the fringe, which is normalized to one.[11]

Firms' Production Technology In each industry, superstar firms and small firms in the fringe produce their variety using a linear production technology in labor:

$$y_{ijt} = q_{ijt} l_{ijt} \quad \text{and} \quad y_{ckjt} = q_{cjt} l_{ckjt} \tag{7}$$

[10] In Section 5, 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 removing combative advertising, although the exact magnitudes change.

[11] We should note that we model advertising as a static decision. The marketing literature typically assumes that advertising has a so-called carry-over effect, i.e., current advertising affects future sales through an advertising stock of goodwill. However, the marketing 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, Hitsch, and Manchanda (2005), Doganoglu and Klapper (2006), Danaher, Bonfrer, and Dhar (2008), Terui, Ban, and Allenby (2011), Shapiro, Hitsch, and Tuchman (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 still acts as a demand shifter and would lead to mechanisms similar to those obtained in our current model regarding its interaction with innovation and misallocation.

where l_{ijt} and l_{ckjt} denote the labor input, 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 level of productivity, which can be built over time through R&D and 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:

$$R_{ijt} = \chi z_{ijt}^{\phi} Y_t \tag{8}$$

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 $n_{ijt}^{k} \in \{-\bar{n}, -\bar{n}+1, ..., \bar{n}-1, \bar{n}\}$ holding:

$$\frac{q_{ijt}}{q_{kjt}} = (1+\lambda)^{n_{ijt}^{k}} \tag{9}$$

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 $q_{jt}^{leader} \equiv \max_{k=1,...,N_{jt}}\{q_{kjt}\}$.[12]

[12]Note that the parameter ζ governs the productivity of the competitive fringe as a whole. An increase in its value increases the market share of all small firms compared to the superstars. The individual market share of any small firm is infinitesimally small compared to any superstar since its Lebesgue measure is zero. Thus, ζ can take any positive value, including above unity, without implying that any single small firm is more productive than any given superstar. Letting $\zeta > 1$ gives our model the flexibility to entertain scenarios in which the competitive fringe as a whole can have a market share larger than any superstar.

Advertising Each superstar firm can spend resources on advertising its product to affect the quality shifter $\hat{\omega}_{ijt}$. In order to achieve a quality shifter ω_{ijt}, firm i of industry j must spend

$$A_{ijt} = \chi_a \omega_{ijt}^{\phi_a} Y_t \tag{10}$$

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 is given by

$$R_{kjt}^{e} = \nu X_{kjt}^{\epsilon} Y_t. \tag{11}$$

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 $X_{jt} = \int X_{kjt} dk = m_{jt} X_{kjt}$. 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 τ. For entry, we assume that there is a measure one of entrepreneurs who pay a cost $\psi e_t^2 Y_t$ 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.

2.2 Equilibrium

Household's Problem Household utility maximization delivers the Euler equation:

$$\frac{\dot{C}_t}{C_t} = r_t - \rho. \tag{12}$$

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:[13]

$$y_{ijt} = \hat{\omega}_{ijt}^{\eta} \left(\frac{p_{ijt}}{\tilde{p}_{sjt}}\right)^{-\eta} \left(\frac{\tilde{p}_{sjt}}{p_{jt}}\right)^{-\gamma} \frac{1}{p_{jt}} Y_t \tag{13}$$

$$\tilde{y}_{cjt} = \left(\frac{\tilde{p}_{cjt}}{p_{jt}}\right)^{-\gamma} \frac{1}{p_{jt}} Y_t \tag{14}$$

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 $\tilde{p}_{sjt} \equiv \left(\sum_{i=1}^{N_{jt}} \hat{\omega}_{ijt}^{\eta} p_{ijt}^{1-\eta}\right)^{\frac{1}{1-\eta}}$ as the ideal price index among the different varieties of the superstars and $p_{jt} \equiv \left(\tilde{p}_{sjt}^{1-\gamma} + \tilde{p}_{cjt}^{1-\gamma}\right)^{\frac{1}{1-\gamma}}$ as the ideal price index of the industry. The relative output between any two superstars i and k of the same industry is:

$$\frac{y_{ijt}}{y_{kjt}} = \left(\frac{\hat{\omega}_{kjt}}{\hat{\omega}_{ijt}} \frac{p_{ijt}}{p_{kjt}}\right)^{-\eta} \tag{15}$$

This makes it apparent that firms can use advertising to shift demand toward their products and thereby increase profits at the expense of their direct competitors.[14] 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 a fringe

[13] See Appendix B.1 for a detailed derivation of the set of all static equilibrium conditions.

[14] Note that, in our model, the price-elasticity of demand is an endogenous object that is affected by the advertising choices of the firms. See equations (B.6) and (C.4) in the appendix.

firm belonging to the same industry is:

$$\frac{y_{ijt}}{\tilde{y}_{cjt}} = \hat{\omega}_{ijt}^{\eta} \left(\frac{p_{ijt}}{\tilde{p}_{sjt}}\right)^{-\eta} \left(\frac{\tilde{p}_{sjt}}{\tilde{p}_{cjt}}\right)^{-\gamma} \tag{16}$$

Finally, the allocation of expenditure across industries is determined by the relative price of the industry to the price of the final good, $\frac{1}{p_{jt}}$. Within-industry expenditures hold:

$$\tilde{p}_{sjt}\tilde{y}_{sjt} = \sum_{i=1}^{N_{jt}} p_{ijt}y_{ijt} \quad \text{and} \quad p_{jt}y_{jt} = \tilde{p}_{sjt}\tilde{y}_{sjt} + \tilde{p}_{cjt}\tilde{y}_{cjt} \tag{17}$$

among superstars alone (excluding the fringe), and including superstars and fringe, respectively.

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

$$\max_{y_{ijt},\omega_{ijt}} \left\{ p_{ijt}y_{ijt} - w_t l_{ijt} - \chi_a \omega_{ijt}^{\phi_a} Y_t \right\} \tag{18}$$

subject to equations (13)-(14) 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 of production, so that the price is $p_{ijt} = M_{ijt}\frac{w_t}{q_{ijt}}$. The equilibrium markup is given by:

$$M_{ijt} = \left[\left(\frac{\eta - 1}{\eta}\right) - \left(\frac{\gamma - 1}{\gamma}\right)\sigma_{ijt} - \left(\frac{\eta - \gamma}{\eta\gamma}\right)\tilde{\sigma}_{ijt} \right]^{-1} \tag{19}$$

In this formula, we have defined:

$$\sigma_{ijt} \equiv \frac{p_{ijt}y_{ijt}}{p_{jt}y_{jt}} \quad \text{and} \quad \tilde{\sigma}_{ijt} \equiv \frac{p_{ijt}y_{ijt}}{\tilde{p}_{sjt}\tilde{y}_{sjt}} \tag{20}$$

as, respectively, the market share of firm i among all firms (superstars and fringe) in its industry, and the market share of the firm among the superstars only. Equation (19) shows that a firm's markup is increasing in both of these market share measures. Importantly, we can write market shares in terms of relative outputs and productivities, as follows:

$$\sigma_{ijt} = \frac{\hat{\omega}_{ijt} \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(\frac{y_{kjt}}{y_{ijt}} \right)^{\frac{\eta-1}{\eta}} \right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{\left(\frac{\tilde{y}_{cjt}}{y_{ijt}} \right)^{\frac{\gamma-1}{\gamma}} + \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(\frac{y_{kjt}}{y_{ijt}} \right)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}} \quad \text{and} \quad \tilde{\sigma}_{ijt} = \frac{\hat{\omega}_{ijt}}{\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt} \left(\frac{y_{kjt}}{y_{ijt}} \right)^{\frac{\eta-1}{\eta}}} \tag{21}$$

In combination with the demand functions derived above, this allows us to obtain the following set of static equilibrium conditions:

$$\left(\frac{y_{ijt}}{y_{kjt}} \right)^{\frac{1}{\eta}} = \frac{q_{ijt}}{q_{kjt}} \frac{\hat{\omega}_{ijt}}{\hat{\omega}_{kjt}} \frac{M_{kjt}}{M_{ijt}}, \quad \forall k \neq i \tag{22a}$$

$$\frac{y_{ijt}}{\tilde{y}_{cjt}} = \frac{q_{ijt}}{q_{cjt}} \frac{\sigma_{ijt}}{\sigma_{cjt}} \frac{1}{M_{ijt}}. \tag{22b}$$

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. The static profits before advertising costs ($\pi_{ijt} = p_{ijt} y_{ijt} - w_t l_{ijt}$) are proportional to the product of the superstar's market share and the Lerner index:

$$\pi_{ijt} = \sigma_{ijt} (1 - M_{ijt}^{-1}) Y_t \tag{23}$$

Advertising Choices As with output choices, a superstar firm internalizes that its advertising decisions affect industry-level prices through their effects on the firm's own market shares relative to other superstars and the fringe, as well as on other firms' 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 B.1 for details), we can write:

$$\frac{\sigma_{ijt}}{1+\omega_{ijt}} \left[\frac{N_{jt} - \hat{\omega}_{ijt}}{N_{jt}} + \frac{\gamma - \eta}{(\eta - 1)\gamma} \left(\tilde{\sigma}_{ijt} - \frac{\hat{\omega}_{ijt}}{N_{jt}} \right) + \frac{\eta}{\eta - 1} \frac{\gamma - 1}{\gamma} \sigma_{ijt} \left(\frac{\hat{\omega}_{ijt}}{N_{jt} \tilde{\sigma}_{ijt}} - 1 \right) \right] = \chi_a \phi_a \omega_{ijt}^{\phi_a - 1} \tag{24}$$

As both markups and taste shifters are functions of market shares, and these are themselves functions of relative outputs, equations (22a), (22b) and (24) comprise a system of $2N_{jt}$ equations and $2N_{jt}$ 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}. Since the model does not admit closed-form solutions for (ω, σ, M), the resulting equilibrium relationship between these variables will be discussed within the context of the estimated set of parameters in Section 3.

We denote 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:

$$l_{ijt} = \frac{\sigma_{ijt}}{w^{rel}_t} M^{-1}_{ijt} \quad \text{and} \quad l_{cjt} = \frac{\sigma_{cjt}}{w^{rel}_t} \tag{25}$$

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, $\int_0^1 (l_{cjt} + \sum_{i=1}^{N_{jt}} l_{ijt})\, dj = 1$, gives us a formula for this relative wage:[15]

$$w^{rel}_t = \int_0^1 \left(\sigma_{cjt} + \sum_{i=1}^{N_{jt}} \sigma_{ijt} M^{-1}_{ijt} \right) dj \tag{26}$$

Superstar Value Function and R&D Decision As we have just seen, static production and advertising decisions, markups, and profits within each industry only depend on the number of superstars and the distribution of their relative productivities. Therefore, the relevant state for a firm i in industry j is given by the vector collecting the number of productivity steps relative to all other superstars in the industry, $\mathbf{n}_{ijt} = \{n^k_{ijt}\}_{k \neq i}$, and the number of superstars in the industry, $N_{jt} = |\mathbf{n}_{ijt}| + 1$. Let us drop time subscripts unless otherwise needed. A superstar firm i chooses an innovation rate z_i to maximize the value of the firm, given by:

[15] Note that the relative wage (which, in this economy, is nothing but the aggregate labor share) may be interpreted as the inverse of the aggregate markup, the latter defined as a sales-weighted harmonic average of firm-level markups.

$$
\begin{aligned}
rV(\mathbf{n}_i, N) = \max_{z_i} \Bigg\{ & \pi^{adv}(\mathbf{n}_i, N) - \chi z_i^{\phi} Y \\
& + z_i \left[V\Big(\mathbf{n}_i \backslash \{n_i^k = \bar{n}\} + \mathbf{1}, N - |\{n_i^k = \bar{n}\}|\Big) - V(\mathbf{n}_i, N) \right] - \sum_{\{k: n_i^k = -\bar{n}\}} z_{kj} V(\mathbf{n}_i, N) \\
& + \sum_{\{k: n_i^k > -\bar{n}\}} z_{kj} \left[V\Big(\mathbf{n}_i \backslash \{n_i^k\} \cup \{n_i^k - 1\} \backslash \{n_i^l = \bar{n} + n_i^k\}, N - |\{n_i^l = \bar{n} + n_i^k\}|\Big) - V(\mathbf{n}_i, N) \right] \\
& + X_j \left[V\Big(\mathbf{n}_i \cup \{\min\{\bar{n}, \bar{n} + \min(\mathbf{n}_i)\}\}, \min(N+1, \bar{N})\Big) - V(\mathbf{n}_i, N) \right] \Bigg\} + \dot{V}(\mathbf{n}_i, N) \qquad (27)
\end{aligned}
$$

In this Hamilton-Jacobi-Bellman equation, the first line is the profit flow from sales net of labor and advertising costs, minus the costs from 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 the 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 $\bar{n}$, the latter firm exits, and the number of superstars in the industry decreases by one. 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 $g > 0$, firm value holds $V(\mathbf{n}_i, N) = v(\mathbf{n}_i, N)Y$ for a time-invariant v, so that $\dot{V}(\mathbf{n}_i, N) = gv(\mathbf{n}_i, N)Y$. Using equation (12), we can write:

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

$$- \sum_{\{k:n_i^k=-\bar{n}\}} z_{kj} v(\mathbf{n}_i, N) + X_j \left[v(\mathbf{n}_i \cup \{\min\{\bar{n}, \bar{n} + \min(\mathbf{n}_i)\}\}, \min(N+1, \bar{N})) - v(\mathbf{n}_i, N) \right] \Bigg\}. \tag{28}$$

The optimal level of innovation is given by:

$$z_i = \left(\frac{v\Big(\mathbf{n}_i \backslash \{n_i^k = \bar{n}\} + \mathbf{1}, N - |\{n_i^k = \bar{n}\}|\Big) - v(\mathbf{n}_i, N)}{\chi \phi} \right)^{\frac{1}{\phi - 1}}. \tag{29}$$

Small Firm Innovation To obtain the optimal behavior of small firms and the entry into the superstar status, we define $\Theta = (N, \vec{n})$ as the state of the industry, where $N \in \{1, ..., \bar{N}\}$ is the number of superstars in the industry and $\vec{n} \in \{0, ..., \bar{n}\}^{N-1}$ 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 Θ (with $N(\Theta) < \bar{N}$), each small firm in the competitive fringe chooses R&D investment to maximize:

$$\begin{aligned} rV^e(\Theta) = \max_{X_{kj}} \Big\{ & X_{kj} V\Big(\{\tilde{\mathbf{n}}_j - \bar{n}\} \cup \{-\bar{n}\}, N_j + 1\Big) - \tau V^e(\Theta) - \nu X_{kj}^{\epsilon} Y \\ & + \sum_{\Theta'} p(\Theta, \Theta') \Big(V^e(\Theta') - V^e(\Theta) \Big) \Big\} + \dot{V}^e(\Theta) \end{aligned} \tag{30}$$

where $V^e(\Theta)$ is the value of a small firm in industry j and $\tilde{\mathbf{n}}_j = \mathbf{n}_{kj}$, where k denotes a productivity leader in industry j.[16] 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:

$$X_{kj} = \left(\frac{v\big(\{\tilde{\mathbf{n}}_j - \bar{n}\} \cup \{-\bar{n}\}, N_j + 1\big)}{\nu \epsilon} \right)^{\frac{1}{\epsilon - 1}} \tag{31}$$

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

$$v^e(\Theta) = \frac{1}{\rho + \tau} \left[\left(1 - \frac{1}{\epsilon}\right) \frac{v(\{\tilde{\mathbf{n}}_j - \bar{n}\} \cup \{-\bar{n}\}, N_j + 1)^{\frac{\epsilon}{\epsilon - 1}}}{(\nu \epsilon)^{\frac{1}{\epsilon - 1}}} + \sum_{\Theta'} p(\Theta, \Theta') \Big(v^e(\Theta') - v^e(\Theta) \Big) \right] \tag{32}$$

[16] Note that we use $\int_{k=i} V_k^e(\Theta) dk = 0$ in the first term, i.e., the value of the small firm is insignificant compared to the value of the superstar firm it becomes, since it is of mass zero in the competitive fringe.

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 mass of industries of type Θ.[17] The value of being an entrepreneur, denoted S, is:

$$\rho S = \max_{e}\{-\psi e^2 Y + eW\} \tag{33}$$

Guessing and verifying that $S = sY$ in a BGP, we obtain that:

$$e = \frac{1}{2\psi}\sum_{\Theta} v^e(\Theta)\mu(\Theta) \tag{34}$$

which implies $s = \frac{1}{4\psi\rho}\left[\sum_{\Theta} v^e(\Theta)\mu(\Theta)\right]^2$. In a stationary equilibrium, entry into the competitive fringe equals exit from the competitive fringe, implying $e = \tau m$. In combination with equation (34), we get the equilibrium measure of small firms in the economy:

$$m = \frac{\sum_{\Theta} v^e(\Theta)\mu(\Theta)}{2\psi\tau} \tag{35}$$

Equilibrium Definition The 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 mass of small firms m_t, and a set of vectors $\{\mathbf{n}_{ijt}\}$ that denote the relative productivity distance between firm i and every other firm in the same industry j at time t, such that, $\forall t \geq 0$, $j \in [0,1]$, $i \in \{1, ..., N_{jt}\}$:

(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.

[17] We can show that the expected value of $\sum_{\Theta'} p(\Theta, \Theta')(V^e(\Theta') - V^e(\Theta))$ in a stationary equilibrium is equal to zero (see Proposition 1 in Appendix B.3). W is thus equal to $\frac{1-\frac{1}{\epsilon}}{\rho+\tau}(\nu\epsilon)^{\frac{1}{1-\epsilon}} \int_0^1 V(\{\tilde{\mathbf{n}}_j - \bar{n}\} \cup \{-\bar{n}\}, N_j + 1)^{\frac{\epsilon}{\epsilon-1}}\, dj$.

(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:

$$Y_t = C_t + \int_0^1 \sum_{i=1}^{N_{jt}} \chi z_{ijt}^{\phi} Y_t \, dj + \int_0^1 \sum_{i=1}^{N_{jt}} \chi_a \omega_{ijt}^{\phi_a} Y_t \, dj + \int_0^1 m_t \nu X_{kjt}^{\epsilon} Y_t \, dj + \psi e_t^2 Y_t \tag{36}$$

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.

Growth Rate Finally, the growth rate of aggregate output in this economy at time t is given by:[18]

$$g_t = -g_{w^{rel},t} + \ln(1+\lambda) \sum_{\Theta} p_{lit}(\Theta)\mu_t(\Theta) + \sum_{\Theta}\sum_{\Theta'} \Big(f_t(\Theta') - f_t(\Theta) \Big) p_t(\Theta, \Theta')\mu_t(\Theta) \tag{37}$$

where $p_{li}(\Theta)$ is defined as the arrival rate of a leader innovation. In this equation, $g_{w^{rel},t}$ is the growth rate of the relative wage $w_t^{rel} \equiv \frac{w_t}{Y_t}$; the second term comes from the growth rate of the industry leaders; and the third term accounts for production reallocation as industries move between states, where we have defined:

$$f_t(\Theta) \equiv \frac{1}{\gamma - 1} \ln \left(1 + \left(\sum_{i=1}^{N_t(\Theta)} \hat{\omega}_{it}(\Theta) \left(\frac{y_{it}}{\tilde{y}_{ct}}(\Theta) \right)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}} \right) \tag{38}$$

In a balanced growth path with a time-invariant distribution over Θ, we have $g_{w^{rel},t} = 0$ and $\mu_t(\Theta) = \mu(\Theta)$. Therefore, the BGP rate of economic growth is given by:

$$g = \ln(1+\lambda) \sum_{\Theta} p_{li}(\Theta)\mu(\Theta) \tag{39}$$

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. The innovation by all other firms affects this growth rate through their influence on the stationary industry state distribution $\mu(\Theta)$ and their

[18] See Appendix B.2 for the full derivation.

strategic effect on leader innovation.

3 Quantitative Analysis

3.1 Estimation

The main focus in our counterfactual exercises in Section 4 will be to understand the static and dynamic implications of the interaction between advertising and innovation both within and across industries, as well as for the aggregate economy. Thus, our estimation strategy requires that the model is consistent with empirical observations regarding advertising, innovation, and market concentration.

In particular, the model is estimated to replicate two within-industry inverted-U shaped relationships observed in the data: (i) an inverted-U relationship between innovation and firms' market share, and (ii) an inverted-U relationship between advertising expenditures and firms' market share.[19] Matching both of these margins helps carefully discipline the implications for innovation, economic growth, and welfare in the various counterfactual exercises of Section 4.

We estimate the model at an annual frequency, and set the consumer discount rate externally to $\rho = 0.04$. This leaves 12 parameters to estimate: the innovation step size, λ; the R&D cost scale parameters for superstars and small firms, (χ, ν); the corresponding R&D cost curvature parameters, (ϕ, ϵ); the relative productivity of the competitive fringe compared to the leader, 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); and two elasticities of substitution: the elasticity among superstars' varieties within an industry, η; and the elasticity between the superstars' combined output and the fringe's combined output, γ. Because of the non-linearities of the model, individual moments cannot uniquely identify each parameter separately. We therefore estimate the 12 parameters jointly through a simulated method of moments (SMM) estimation procedure. The identification success of this method requires that we choose moments which are sufficiently sensitive to variations in the structural parameters. We describe these moments next, and relegate to Appendix A.2 all the

[19] See Cavenaile, Celik, and Tian (2021) for the documentation of both regularities, and their robustness across different specifications.

details regarding the data sources and the way these moments are computed. For a discussion of which moments help identify which parameters, and the Jacobian matrix of the model moments with respect to each estimated parameter, see Appendix A.3.

TABLE 1: BENCHMARK MODEL PARAMETERS AND TARGET MOMENTS

A. Parameter estimates

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
β(innovation, relative sales)	0.629	0.982
Top point (innovation, relative sales)	0.505	0.483
β(advertising, relative sales)	6.260	7.614
Top point (advertising, relative sales)	0.533	0.521

Notes: The estimation is done with the Simulated Method of Moments. Panel A reports the estimated parameters. Panel B reports the simulated and empirical moments.

Table 1 presents the results of our SMM estimation exercise for the United States. Panel A reports the parameter values, and Panel B reports the results in terms of moment matching. The model manages to match the data moments well despite overidentification. We target a combination of aggregate and industry-level moments. At the aggregate level, we target the growth rate of real GDP per capita, the aggregate R&D and advertising expenditures over GDP, the sales-weighted average and standard deviation of firm-level markups, the labor share, the firm entry rate, the average firm profitability, the average relative quality of the leader, and its standard deviation across industries.

The remaining four moments pertain to the indirect inference exercise which helps the model reproduce the two non-linear relationships observed in the data. To discipline the inverted-U shaped relationship between innovation and firms' market share within industries, we target the linear term and the top point of the inverted-U, which we obtain from the coefficients of an intra-industry regression of a firm's innovation on its relative sales and the square of relative sales. Likewise, to ensure that the model can reproduce the inverted-U shaped intra-industry relationship between advertising and firm market share observed in the data, we target the corresponding linear term and the top point from an intra-industry quadratic regression of firm advertising expenditures on the firm's relative sales and relative sales squared.

3.2 Optimal Advertising and Innovation Policies

Using the estimated set of parameters presented above, Figure 1 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).[20] 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 E.1 in the Appendix shows the corresponding policy functions for innovation, exhibiting a similar feature: firms innovate the most when they are close

[20]Since we assume that there exists a maximum technology gap $\bar{n}$ between any two superstars, some states on the right panel of Figure 1 are illegal, which is why the policy function is displayed as a strip in the space of states.

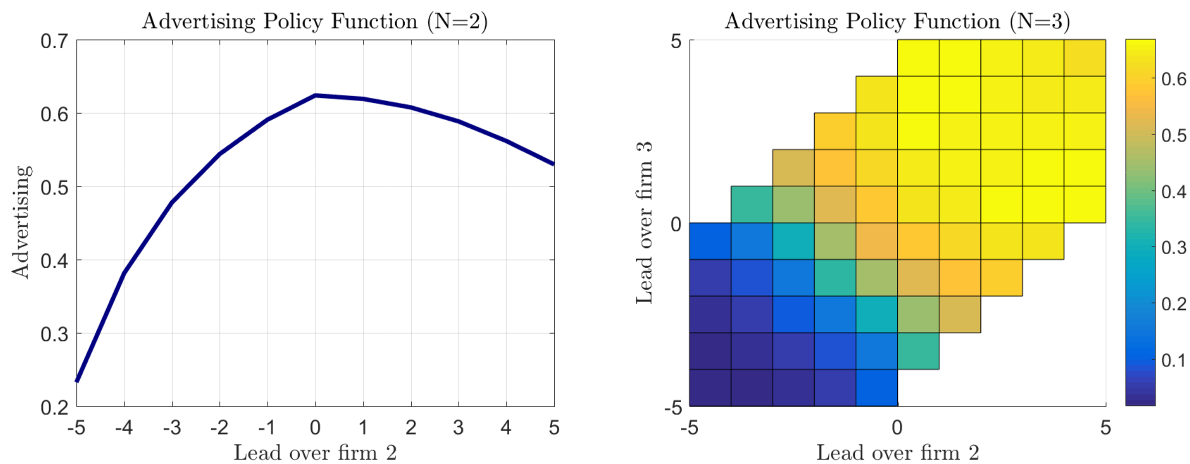

FIGURE 1: ADVERTISING POLICY FUNCTION

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).

to being neck-to-neck, and innovation incentives decrease the higher the technological gap with their competitors.

Our estimated economy also exhibits firm-level correlation patterns between markups, R&D expenditures, and advertising expenditures consistent with those found in the data (e.g., De Loecker, Eeckhout, and Unger (2020)), which are shown in Table E.8. All are all positively correlated at the firm level, consistent with the idea that firms use both types of intangibles to increase their profits and harness greater market power.

3.3 Advertising and Innovation Within and Across Industries

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 relationship between innovation, advertising, and competition.

Figure 2 shows that the estimated model is able to replicate the inverted-U shaped relationship between innovation and market share, and between advertising and market share (recall that the intercept and top point of both of these curves were targets of the estimation). The figure displays firm-level R&D (left panel) and advertising (right panel) expenditures in the model as functions of

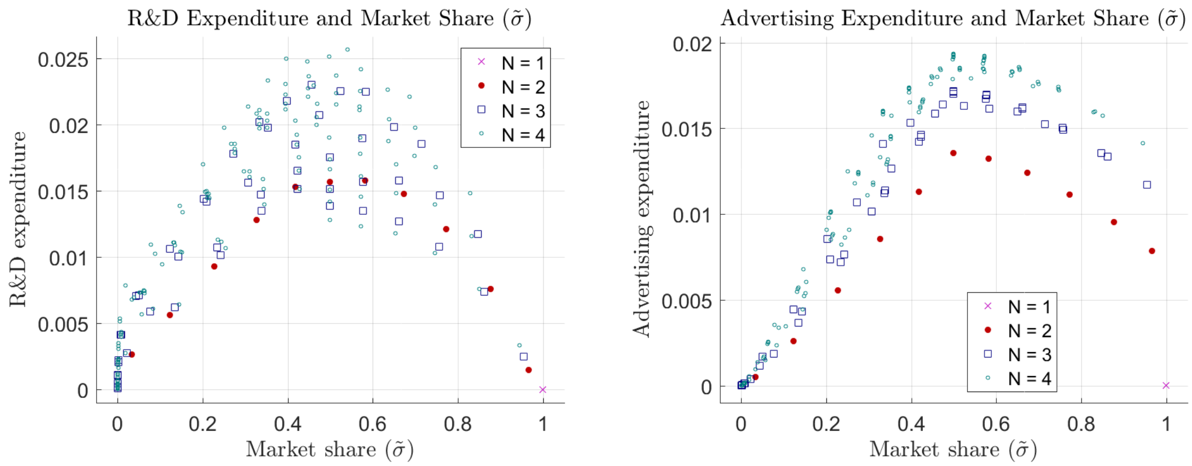

FIGURE 2: R&D EXPENSES, ADVERTISING, AND FIRM MARKET SHARES

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 equation (20)). 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 firm's market share relative to other superstars in its industry (i.e., $\tilde{\sigma}$ defined in equation (20)).[21] 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. Note that the inverted-U relationships continue to hold even within industries with the same number of superstars N, which is also true in the data.

4 Counterfactual Experiments

How does advertising affect the macroeconomy? How does it affect social welfare, and what are the implications for government intervention? In this section, we perform counterfactual experiments to study how advertising and its interaction with R&D affects macroeconomic aggregates such as the average markup and its dispersion, the labor share, and long-run economic growth. We also study the welfare implications of advertising in the short and long run. We investigate policy implications of our model by considering the linear taxation/subsidization of advertising. Finally, we consider

[21]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 U.S. 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.

the heterogeneous firm-level effects of the advertising shutdown and optimal advertising taxation counterfactuals to identify their winners and losers.

4.1 The Macroeconomic Effects of Shutting Down Advertising

As a first pass to analyzing the macroeconomic effects of advertising, we conduct a counterfactual experiment in which we shut down advertising completely. In particular, we study how our quantitative results change if superstar firms are not able to invest in advertising.[22] In this case, the perceived quality of every single variety is equal to one. We analyze how shutting down advertising affects macroeconomic aggregates, static allocative efficiency, and welfare, compared to our baseline estimated economy.

4.1.1 The Dynamic Impact on Macroeconomic Aggregates

Table 2 reports the results from our experiment for macroeconomic aggregates. We can first notice that R&D intensity 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 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. Therefore, whether innovation and advertising are substitutes or complements in general equilibrium is theoretically indeterminate, and quantification is needed to reach a conclusion.

Our estimation results in Table 2 suggest that the second effect dominates in the United States, and that R&D and advertising are substitutes at the aggregate level in general equilibrium, since innovation by superstars increases in response to shutting down advertising.[23] This result is in line

[22] This experiment is equivalent to the limiting case of our model in which the cost scale parameter of advertising χ_a goes to infinity.

[23] Note that, 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.

with the empirical findings in Cavenaile and Roldan-Blanco (2021), providing an 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 4.1.2, 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. 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 away from their competitors, and charge higher markups. As a result, shutting down advertising leads to a significant decrease in markups. The average net markup decreases from 0.34 to 0.25. In other words, advertising is found to be responsible for roughly one quarter of the average net markup observed in the estimated 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. Our findings highlight the importance of accounting for advertising in the determination of markups, which would be attributed to other channels if one were to ignore advertising as a potential mechanism to influence the demand for a firm's output.

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 sound beneficial for social welfare at first glance, the effects of shutting down advertising on static and

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.

dynamic allocative efficiency are found to be quite nuanced, which we investigate next.

4.1.2 The Impact on Static Misallocation of Resources and Competition

By acting as a demand shifter, advertising by superstars can affect the relative production of different firms within each industry. In our model, larger firms charge higher markups which creates static misallocation: more productive firms do not demand enough labor and produce too little relative to the efficient allocation. As a result, heterogeneous advertising, by shifting market shares between incumbents, could directly affect the degree of static allocative efficiency.

In the previous section, we had established that shutting down advertising led to a significant decrease in the average net markup and its dispersion by one quarter. Therefore, one might be tempted to expect an increase in allocative efficiency. Surprisingly, we find the opposite result to be the case: shutting down advertising reduces allocative efficiency, decreasing the level of output by 4.63% of its value.

This result owes to two effects 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}). This means that, statically, the economy with advertising allocates inputs more efficiently, even if we hold perceived qualities $\{\hat{\omega}_{ij}\}_{i=1}^{N_j}$ fixed. Second, optimal advertising chosen by the superstars in equilibrium is such that the perceived quality of large and more efficient superstars is magnified compared to the smaller and less efficient superstars. This further amplifies the gains from production by improving the perceived quality of the more abundant (and cheaper to produce) varieties. 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.

In a typical industry, the optimal level of advertising tends to both shift production from smallest to largest firms and to raise the relative quality of more productive firms. In addition, advertising depresses relative wages. All these forces improve allocative efficiency and increase industry output in our baseline calibration compared to our counterfactual economy without advertising.

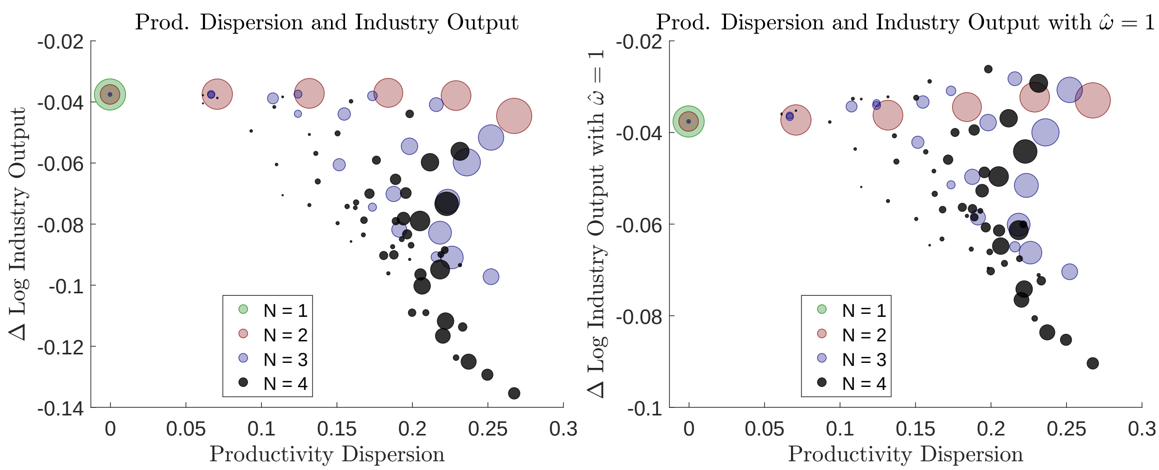

FIGURE 3: CHANGE IN INDUSTRY OUTPUT BY PRODUCTIVITY DISPERSION

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 presents 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.

Focusing on how shutting down advertising affects allocative efficiency across different industry states also reveals interesting patterns. The left panel of Figure 3 displays the change in industry output as a function of productivity dispersion in the industry. As discussed above, these changes are the combined result of changes in the labor allocation between firms with different productivities and the change in their perceived quality ($\hat{\omega}$). The right panel of Figure 3 shows how much of the change in industry output is due to labor reallocation alone. In particular, it shows the change in industry output evaluated at a fixed perceived quality, $\hat{\omega}_{ij} = 1$.[24] As a result, the change in industry output in that panel is totally due to changes in quantities produced. Comparing both panels of Figure 3 shows that labor reallocation between firms explains a large share of the change in industry output and misallocation. We can also notice that the reallocation of market shares, improving static allocative efficiency, is stronger in industries where the dispersion in terms of productivity is larger.

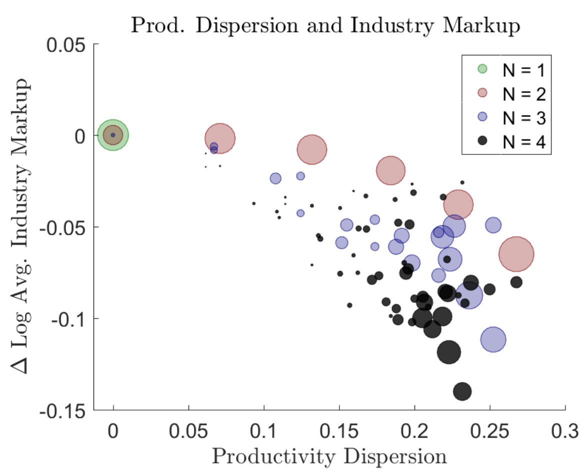

FIGURE 4: CHANGE IN MARKUP BY PRODUCTIVITY DISPERSION

Notes: This figure displays the changes in the average industry markup as a function of productivity dispersion. Each circle presents 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.

Figure 4 further displays the reduction in the average industry markup once again as a function of productivity dispersion. One can see that the decline of the average industry markup tends to be stronger in industries with higher productivity dispersion. However, despite its positive effects, the reallocation and perceived quality amplification channels dominate, as we observe an overall decline in industry output across the board for all industry states (Figure 3).

A comparable decomposition is reported in Figure 5 which depicts the difference in industry

[24] In the absence of advertising, all $\hat{\omega}_{ij}$ are equal to one.

output between our baseline model and our counterfactual economy as a function of industry concentration measured by the Herfindahl-Hirschman Index (HHI), taken from the baseline economy. Once again, the left panel shows the full change in industry output and the right panel evaluates industry output when $\hat{\omega}_{ij}$ are set to one. It confirms that the decrease in industry output is larger for industries that are more concentrated and that labor reallocation across firms plays a significant role in explaining changes in static misallocation.

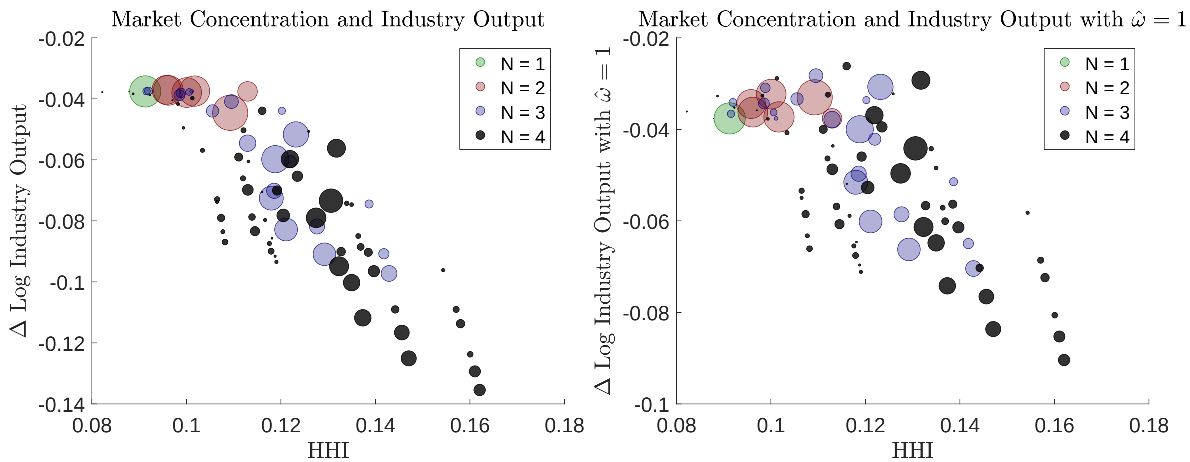

FIGURE 5: CHANGE IN INDUSTRY OUTPUT BY MARKET CONCENTRATION

Notes: This figure depicts the difference in industry output between the baseline model and the counterfactual economy without advertising as a function of industry concentration measured by the HHI (taken from the baseline economy). Each circle presents 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.

4.1.3 Short-Run versus Long-Run Effects on Markups and Welfare

In this section, we investigate the short- and long-run effects of shutting down advertising on markups and social welfare. First, we decompose our main results regarding markups between the static and dynamic parts. The static part results from changes in markups due to shutting down advertising 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. This leads to a change in the distribution over industry states with different markups. Second, we study the welfare implications of shutting down advertising, and perform a similar decomposition between the short-run and long-run welfare changes.[25] Shutting down advertising changes industry output and static allocative efficiency as shown in Section 4.1.2. In addition, it also affects R&D investment and hence both the stationary distribution over industry states and the growth rate of the economy.

First, we can decompose the change in markups between our baseline calibration and our counterfactual economy into a static and dynamic effect. Statically, advertising affects the markups that superstar firms charge as well as the distribution of market shares within industry. The change in aggregate markups between the two economies that results from those changes for a fixed distribution over industry states is what we call the static effect of advertising on markups. The dynamic effect arises from the impact of changes on advertising on the R&D investment of superstar firms, which further leads to a change in the distribution over industry states. This dynamic effect is the result of the equilibrium interaction between advertising and R&D. Statically, we find that shutting down advertising reduces the average net markup by 22.3% from 0.342 to 0.266. The dynamic effect coming from the interaction between advertising and R&D investment and its impact on the industry state distribution further reduces average net markup by 4.59% to 0.254. At the same time, the dispersion of markups also goes down when advertising is shut down (from 0.44 to 0.34). Around three quarters of this decrease is due to the static effect of advertising, whereas the remainder owes to the long-run change in the stationary distribution across industry states.

Second, regarding welfare, our model allows for an analytical decomposition of the change in welfare (W) between our baseline calibration and our counterfactual economy without advertising, as follows (see the details in Appendix B.4):

$$\Delta W = \frac{1}{\rho}\left[-\Delta \ln w^{rel} + \Delta \sum_{\Theta} f(\Theta)\mu(\Theta) + \Delta \ln\left(\frac{C}{Y}\right)\right] + \frac{1}{\rho^2}\Delta g \tag{40}$$

The first term in 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 output level, Y_0. 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.

[25] In Appendix D, we state the full social planner's problem and derive closed-form solutions for the static part.

TABLE 3: ADVERTISING SHUTDOWN: SHORT-RUN VS. LONG-RUN EFFECTS ON EFFICIENCY

	Static		Static+New Distribution		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 markups and social welfare between our baseline calibration and our counterfactual economy without advertising. The first two columns in the table report the static effect of advertising on markup and social welfare, i.e., fixing the distribution over industry states and the level of R&D. The third and fourth columns in this table display what happens if we further let the distribution adjust (but still keep R&D and growth fixed). The last two columns show the full results including the dynamic effects due to changes in R&D investment.

Table 3 shows how each of these components is affected by shutting down advertising. Overall, we obtain a welfare loss of 0.86% in consumption-equivalent terms.[26] The first two columns in the table report the static effect of advertising on welfare, i.e., fixing the distribution over industry states and the level of R&D. 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 4.1.2, 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, total output of superstars increases which results in welfare gains. On the other hand, the relative wage further decreases. Finally, the last two columns of Table 3 show the full results including the dynamic effects due to changes in R&D investment.[27] Shutting down advertising raises the consumption-to-output ratio as a result of changes in total R&D and advertising expenses. In

[26]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 (B.17) of Appendix B.4.

[27]The welfare numbers are calculated by comparing the two stationary equilibria. We can also feasibly solve for the non-stationary equilibrium 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. 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.

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. Therefore, we conclude 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.

4.2 Should We Tax or Subsidize Advertising?

Our results so far raise some questions in terms of policy implications, especially regarding advertising. Our results from Section 4.1.3 show that totally shutting down advertising is not socially desirable as it reduces welfare. In this section, we study whether a subsidy or a tax on advertising could be welfare-improving. In particular, 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.

Table 4 reports the results of our policy experiment for different values of taxes and subsidies.[28] 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. That is, the substitution effect between advertising and innovation dominates. Taxing advertising also results in a decrease in 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.

Overall, we have several forces associated with taxation or subsidization of advertising that go in opposite directions regarding welfare. We find that there exists an optimal level of tax on advertising that maximizes welfare equal to 62.9% (see Figure 6). This tax is associated to 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

[28] 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.

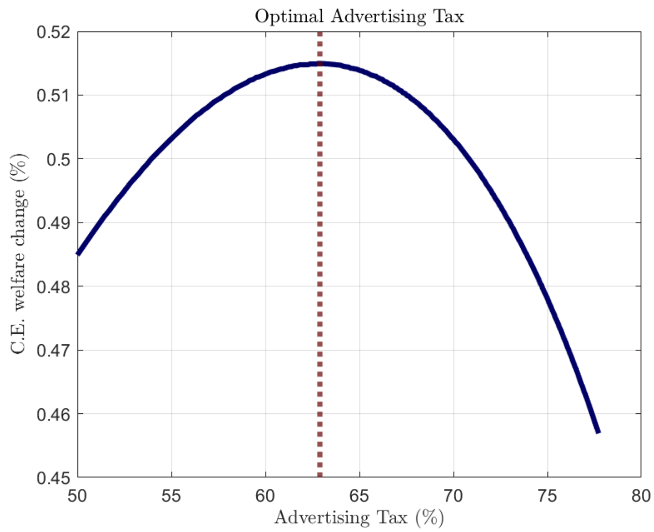

FIGURE 6: THE DYNAMIC WELFARE IMPACT OF ADVERTISING TAXES

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

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.

Section 4.1.3 established that shutting down advertising improved welfare. How does one reconcile this finding with the fact that the optimal linear tax on advertising is found to be positive, and indeed quite high at 62.9%? 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 consumption-to-output ratio, 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.

Table 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.

In most advanced economies including the United States, advertising expenses are not taxed.[29] Our quantitative analysis demonstrates that advertising is a useful activity insofar as it improves static allocative efficiency through a reduction 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 the GDP under the optimal tax rate – while simultaneously improving dynamic efficiency.[30] While the optimal level calculated at 62.9% may seem rather high, this is well within the range European countries levy on petroleum products, which create a large dead-weight loss as well as increase transportation costs.[31] 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.

4.3 The Heterogeneous Effects of Advertising Shutdown and Taxes 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 estimated baseline economy.

The top two panels of Figure E.2 in the Appendix 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.[32] The bottom

[29] 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. The rates that are being considered are quite low in comparison – e.g., 3% in Canada.

[30] 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.

[31] 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 4 shows, a modest 25% tax rate can still deliver 58.3% of the consumption-equivalent welfare gains the optimal tax rate of 62.9% provides.

[32] While the figures for 4-superstar industries are not shown due to space constraints, they are very similar.

two panels of Figure E.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 E.3 repeat the same exercise for market shares instead of firm value.

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 from shutting down advertising and 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.

5 Extensions

One might reasonably wonder whether our quantitative findings are contingent on the specific assumptions regarding how advertising enters consumer preferences, or the modeling choices regarding how the advertising efforts by firms affect the demand shifters (perceived quality $\hat{\omega}$) in the same industry. Motivated by such concerns, in this section, we propose two extensions to our model

– deceptive advertising and non-combative advertising – which can be interpreted as putting more weight on the persuasive and informative views of advertising compared to the baseline, respectively. We further test the robustness of our results by considering an alternative model in which static product market competition is in prices *à la* Bertrand rather than in quantities.

Repeating the quantitative experiments under these extended models reveals that, even under the most extreme parametrizations, almost all of our main findings are preserved, such as the optimality of positive advertising taxes, the aggregate substitution between innovation and advertising, implications on business dynamism, the positive role of advertising in reducing static misallocation in oligopolistic markets, and its overall usefulness for raising welfare, demonstrating the robustness of our main conclusions.

5.1 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 for instance a discussion in 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 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" due to the crowding-out effect.

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 the same way 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. In particular, we assume that, at every instant, advertising in any industry turns out to be (unexpectedly) purely manipulative with probability $\delta \in [0, 1]$.[33] 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.

Appendix C.1 derives the closed-form expressions for welfare this alternative model implies. 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.

TABLE 5: 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.

First, we repeat the advertising shutdown experiment for different values of $\delta \in \{0, 0.25, 0.50, 0.75, 1\}$. Table 5 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 hurt welfare.

Next, we calculate the welfare-maximizing advertising taxes and the associated positive changes

[33] 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.

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 E.1 summarizes the changes to the economy. Interestingly, the optimal tax rate, while higher, 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 5. However, a benevolent government would still choose to tax advertising at a high rate rather than 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.

5.2 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.[34] Therefore, our combative advertising assumption in the baseline model might be too severe, and could be driving our results.

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

$$\hat{\omega}_{ijt} \equiv \frac{1+\omega_{ijt}}{\Lambda + \frac{1-\Lambda}{N_{jt}} \sum_{k=1}^{N_{jt}} (1+\omega_{kjt})} \tag{41}$$

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.

Appendix C.2 derives the equilibrium conditions of this alternative model. 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-estimated. To prove the robustness of our results, we pick the extreme value of $\Lambda = 1$ as opposed to our baseline's $\Lambda = 0$, re-estimate the model, and repeat the experiments. Table E.2 presents the estimated parameter values and the details of the SMM estimation, whereas Table E.3 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

[34] For instance, Cavenaile, Celik, Perla, and Roldan-Blanco (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.

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.

5.3 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, estimate it, and repeat the quantitative experiments. This reveals that almost all of our results are maintained.

Changing the assumption regarding competition alters the static strategic choices that influence the market share and advertising distributions in each industry, holding the industry state Θ fixed. This, in turn, changes the static profit flows accruing to firms under different states of the industry, and thereby influences the dynamic innovation choices of all firms and the business creation decision of the entrepreneurs, and consequently the endogenous market structure through the changes in the stationary industry state distribution $\mu(\Theta)$. Appendix C.3 provides the details of how the best responses of firms and the static equilibrium conditions in each industry change.

We estimate the parameters of this alternative model using the same methodology as the baseline analysis. Table E.4 presents the estimated parameter values and the model fit. Using this estimated model, we repeat the counterfactual experiments, the results of which are displayed in Table E.5. 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-estimating 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, Carvalho, and Grassi (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.

6 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 the 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 proposed a unified approach 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 market share distribution), markups, and growth are all endogenous. In the model, large firms make production, innovation, and advertising choices strategically in an oligopolistic environment, and small firms spend resources on R&D to join the pool of large firms. In equilibrium, large firms of different productivities use advertising strategically to gain market share and charge higher markups, which has implications for allocative efficiency. Dynamically, advertising and innovation choices interact and, at the aggregate level, affect the pace of economic growth.

We estimate the model to match features of the data at different levels of aggregation, and particularly to fit the existing non-linear relationship between innovation, advertising, and competition in the data. We find that advertising has important quantitative implications for macroeconomic aggregates. Since advertising and innovation are substitutes in the estimated model, shutting down advertising improves the incentives of firms to innovate instead, which boosts economic growth. This substitution effect is in line with the empirical findings in Cavenaile and Roldan-Blanco (2021). All in all, while the average net markup decreases by one quarter of its value relative to the baseline estimation when advertising is shut down, the rate of economic growth increases by about 3%. 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.

We then ask whether advertising should be subsidized or taxed on the basis of our estimated model. We find that advertising should be taxed, and that the optimal advertising tax (that is, the tax that maximizes long-run social welfare) 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. Our positive taxation result is also found to be robust across all our model extensions. It therefore provides a justification for the recent efforts by policymakers to impose taxes on (digital) advertising. We believe that these results are relevant for industrial policy as well as public finance, and expect future research to delve further into these and other topics relating the interaction between firm intangibles to competition, growth, and social welfare.

References

AFROUZI, H., A. DRENIK, AND R. KIM (2021): "Growing by the Masses-Revisiting the Link between Firm Size and Market Power," *Working Paper*.

AGHION, P., A. BERGEAUD, T. BOPPART, P. J. KLENOW, AND H. LI (2022): "A Theory of Falling Growth and Rising Rents," *The Review of Economic Studies, forthcoming*.

AGHION, P., C. HARRIS, P. HOWITT, AND J. VICKERS (2001): "Competition, Imitation and Growth with Step-By-Step Innovation," *The Review of Economic Studies*, 68(3), 467–492.

AKCIGIT, U., AND S. T. ATES (2022): "What Happened to US Business Dynamism?," *Journal of Political Economy, forthcoming*.

AKERLOF, G. A., AND R. J. SHILLER (2015): "Phishing for Phools," in *Phishing for Phools*. Princeton University Press.

ARGENTE, D., D. FITZGERALD, S. MOREIRA, AND A. PRIOLO (2021): "How Do Firms Build Market Share?," *Working Paper*.

ARKOLAKIS, C. (2010): "Market Penetration Costs and the New Consumer Margin in International Trade," *Journal of Political Economy*, 118(6), 1151–1199.

ATKESON, A., AND A. BURSTEIN (2008): "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98(5), 1998 – 2031.

BAGWELL, K. (2007): "The Economic Analysis of Advertising," *Handbook of Industrial Organization*, 3, 1701 – 1844.

BARKAI, S. (2020): "Declining Labor and Capital Shares," *The Journal of Finance*, 75(5), 2421–2463.

BASLANDZE, S., J. GREENWOOD, R. MARTO, AND S. MOREIRA (2023): "The Expansion of Product Varieties in the New Age of Advertising," *Working Paper*.

BECKER, G. S., AND K. M. MURPHY (1993): "A Simple Theory of Advertising as a Good or Bad," *The Quarterly Journal of Economics*, 108(4), 941 – 964.

BELO, F., X. LIN, AND A. M. VITORINO (2014): "Brand Capital and Firm Value," *Review of Economic Dynamics*, 17(1), 150 – 169.

BENHABIB, J., AND A. BISIN (2002): "Advertising, Mass Consumption and Capitalism," *Working Paper, New York University*.

——— (2011): "Social Construction of Preferences: Advertising," *Handbook of Social Economics (eds. Jess Benhabib, Alberto Bisin and Matt Jackson)*, pp. 201 – 220.

BURSTEIN, A., V. M. CARVALHO, AND B. GRASSI (2020): "Bottom-up Markup Fluctuations," *NBER Working Paper 27958*.

BUTTERS, G. R. (1977): "Equilibrium Distributions of Sales and Advertising Prices," *The Review of Economic Studies*, 44(3), 465 – 491.

CAVENAILE, L., M. A. CELIK, J. PERLA, AND P. ROLDAN-BLANCO (2023): "A Model of Product Awareness and Industry Life Cycles," *Working Paper*.

CAVENAILE, L., M. A. CELIK, AND X. TIAN (2021): "Are Markups Too High? Competition, Strategic Innovation, and Industry Dynamics," *Working Paper*.

CAVENAILE, L., AND P. ROLDAN-BLANCO (2021): "Advertising, Innovation, and Economic Growth," *American Economic Journal: Macroeconomics*, 13(3), 251–303.

CELIK, M. A., X. TIAN, AND W. WANG (2022): "Acquiring Innovation under Information Frictions," *The Review of Financial Studies*, 35(10), 4474–4517.

CORHAY, A., H. KUNG, AND L. SCHMID (2020a): "Competition, Markups, and Predictable Returns," *The Review of Financial Studies*, 33(12), 5906–5939.

——— (2020b): "Q: Risk, Rents, or Growth?," *USC Marshall School of Business Research Paper*.

COVARRUBIAS, M., G. GUTIÉRREZ, AND T. PHILIPPON (2020): "From Good to Bad Concentration? US Industries over the past 30 years," *NBER Macroeconomics Annual*, 34(1), 1–46.

CROUZET, N., AND J. C. EBERLY (2019): "Understanding Weak Capital Investment: The Role of Market Concentration and Intangibles," *NBER Working Paper 25869*.

DANAHER, P. J., A. BONFRER, AND S. DHAR (2008): "The Effect of Competitive Advertising Interference on Sales for Packaged Goods," *Journal of Marketing Research*, 45(2), 211–225.

DAVID, J. M. (2021): "The Aggregate Implications of Mergers and Acquisitions," *The Review of Economic Studies*, 88(4), 1796–1830.

DAVID, J. M., L. SCHMID, AND D. ZEKE (2022): "Risk-adjusted Capital Allocation and Misallocation," *Journal of Financial Economics*, 145(3), 684–705.

DE LOECKER, J., J. EECKHOUT, AND G. UNGER (2020): "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics*, 135(2), 561 – 644.

DE RIDDER, M. (2020): "Market Power and Innovation in the Intangible Economy," *Working Paper*.

DINLERSOZ, E., N. GOLDSCHLAG, M. YORUKOGLU, AND N. ZOLAS (2023): "On The Role of Trademarks: From Micro Evidence to Macro Outcomes," *Working Paper*.

DINLERSOZ, E. M., AND M. YORUKOGLU (2012): "Information and Industry Dynamics," *American Economic Review*, 102(2), 884 – 913.

DIXIT, A., AND V. NORMAN (1978): "Advertising and Welfare," *The Bell Journal of Economics*, 9(1), 1 – 17.

DOGANOGLU, T., AND D. KLAPPER (2006): "Goodwill and Dynamic Advertising Strategies," *Quantitative Marketing and Economics*, 4(1), 5–29.

DORFMAN, R., AND P. O. STEINER (1954): "Optimal Advertising and Optimal Quality," *American Economic Review*, 44(5), 826 – 836.

DOU, W. W., Y. JI, AND W. WU (2022): "The Oligopoly Lucas Tree," *The Review of Financial Studies*, 35(8), 3867–3921.

DROZD, L. A., AND J. B. NOSAL (2012): "Understanding International Prices: Customers as Capital," *American Economic Review*, 102(1), 364 – 395.

DUBÉ, J.-P., G. J. HITSCH, AND P. MANCHANDA (2005): "An Empirical Model of Advertising Dynamics," *Quantitative Marketing and Economics*, 3(2), 107 – 144.

EINAV, L., P. J. KLENOW, J. D. LEVIN, AND R. MURCIANO-GOROFF (2022): "Customers and Retail Growth," *Working Paper*.

ERICSON, R., AND A. PAKES (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 62(1), 53–82.

FEIJOO MOREIRA, S. (2021): "Provider-driven Complementarity and Firm Dynamics," *Working Paper*.

FITZGERALD, D., S. HALLER, AND Y. YEDID-LEVI (2022): "How Exporters Grow," *Working Paper*.

GLOVER, B., AND O. LEVINE (2017): "Idiosyncratic Risk and the Manager," *Journal of Financial Economics*, 126(2), 320–341.

GOURIO, F., AND L. RUDANKO (2014): "Customer Capital," *The Review of Economic Studies*, 81(3), 1102 – 1136.

GREENWOOD, J., Y. MA, AND M. YORUKOGLU (2022): "'You Will': A Macroeconomic Analysis of Digital Advertising," *Working paper*.

GUTIÉRREZ, G., C. JONES, AND T. PHILIPPON (2021): "Entry Costs and Aggregate Dynamics," *Journal of Monetary Economics*, 124, S77–S91.

HALL, B. H., A. B. JAFFE, AND M. TRAJTENBERG (2001): "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," *NBER Working Paper No. 8498*.

HENNESSY, C., AND T. WHITED (2005): "Debt Dynamics," *Journal of Finance*, (1), 1129–1165.

IGNASZAK, M., AND P. SEDLÁČEK (2022): "Customer Acquisition, Business Dynamism and Aggregate Growth," *Working Paper*.

JERMANN, U., AND V. QUADRINI (2012): "Macroeconomic Effects of Financial Shocks," *American Economic Review*, 102(1), 238–71.

KARABARBOUNIS, L., AND B. NEIMAN (2013): "The Global Decline of the Labor Share," *The Quarterly Journal of Economics*, 129(1), 61–103.

KLEIN, M. A., AND F. ŞENER (2023): "Product Innovation, Diffusion and Endogenous Growth," *Review of Economic Dynamics*, 48, 178–201.

LEONE, R. P. (1995): "Generalizing What is Known about Temporal Aggregation and Advertising Carryover," *Marketing Science*, 14(3), 141–150.

LIND, J. T., AND H. MEHLUM (2010): "With or Without U? The Appropriate Test for a U-Shaped Relationship: Practitioners' Corner," *Oxford Bulletin of Economics and Statistics*, 72(1), 109–118.

MOLINARI, B., AND F. TURINO (2017): "Advertising and Aggregate Consumption: A Bayesian DSGE Assessment," *The Economic Journal*, 128(613), 2106 – 2130.

MORLACCO, M., AND D. ZEKE (2021): "Monetary policy, Customer Capital, and Market Power," *Journal of Monetary Economics*, 121, 116–134.

NIKOLOV, B., AND T. WHITED (2014): "Agency Conflicts and Cash: Estimates from a Dynamic Model," *Journal of Finance*, 69(5), 1883–1921.

PEARCE, J., AND L. WU (2022): "Brand Reallocation, Concentration, and Growth," *SSRN Working Paper*.

RACHEL, L. (2022): "Leisure-Enhancing Technological Change," *Working Paper*.

ROLDAN-BLANCO, P., AND S. GILBUKH (2021): "Firm Dynamics and Pricing under Customer Capital Accumulation," *Journal of Monetary Economics*, 118, 99–119.

SHAPIRO, B. T., G. J. HITSCH, AND A. E. TUCHMAN (2021): "TV Advertising Effectiveness and Profitability: Generalizable Results from 288 Brands," *Econometrica*, 89(4), 1855–1879.

TAYLOR, L. (2010): "Why Are CEOs Rarely Fired? Evidence from Structural Estimation," *Journal of Finance*, 65(6), 2051–2087.

TERRY, S. (2023): "The Macro Impact of Short-Termism," *Econometrica*, forthcoming.

TERUI, N., M. BAN, AND G. M. ALLENBY (2011): "The Effect of Media Advertising on Brand Consideration and Choice," *Marketing Science*, 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

A Estimation Details

A.1 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 estimated via a Simulated Method of Moments (SMM) estimation procedure to match 14 moments in the data. The SMM estimator is defined as the solution to the minimization of the weighted average distance between data and model moments.

In the estimation, we set the maximum number of superstars in an industry to $\bar{N} = 4$ and the maximum productivity step size to $\bar{n} = 5$, which delivers 84 unique industry states Θ. The results do not significantly change if we increase $\bar{n}$ or $\bar{N}$. The estimated value of λ adjusts to absorb the choice of a different $\bar{n}$. The relative productivity of the competitive fringe ζ adjusts to absorb the changes in $\bar{N}$. In the estimated model, $\bar{n}$ is chosen large enough such that the largest superstars that we stop keeping track of are significantly smaller than the leader in terms of revenue and profits in all industry states. Keeping track of these firms would not noticeably change the results.

A.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.[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 averages in our sample.

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

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, Eeckhout, and Unger (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

[1] 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, Jaffe, and Trajtenberg (2001)), we have opted not to include data from the last two years. This decision allows patents sufficient time to accumulate citations.

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

across all years for our sample. In our baseline model, there is no capital. Therefore, the model-generated labor share $w^{rel}L = wL/Y$ corresponds to 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 (2022).

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.

6. **Relationship between firm innovation and relative sales:** Replicating the observed inverted-U relationship between competition and innovation helps us discipline the counterfactual implications of the model regarding economic growth and social welfare. To achieve this, we target the relationship between firm innovation and relative sales. Innovation in the model is measured as the Poisson arrival event of quality improvement, whereas it is measured as average patent citations for each firm in the data.[3] We normalize both by subtracting their means and dividing by their standard deviation. In the data, we regress average citations on relative sales of the firm in its SIC4 industry and their square.[4] The control variables include 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. We target the linear and quadratic terms of a regression of (standardized) average citations on relative sales. The regression results are reported in Table E.6.

7. **Average profitability:** In the model, average profitability is calculated as 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.[5]

10. **Relationship between firm advertising and relative sales:** Replicating the observed inverted-U relationship between competition and advertising also helps use discipline the counterfactual

[3]To measure patent citation, we use the patent grant data obtained from NBER Patent Database Project which covers the years 1976-2006. We 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. We measure innovation as the number of citations a patent received as of 2006. We use the truncation correction weights devised by Hall, Jaffe, and Trajtenberg (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).

[4]The inverted-U result and the quantitative magnitude of the top point 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).

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

implications of the model regarding advertising, economic growth, and social welfare. Therefore, we also require the model-generated relationship between firm advertising expenses and relative sales to be the same as in the data. We regress the logarithm of firm advertising expenses on relative sales of the firm in its SIC4 industry and their square, where xad stands for the advertising expenses found in Compustat.[6] The control variables include 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. We normalize the advertising expenses in both the data and model by subtracting their means and dividing by their standard deviation, and target the coefficients for the linear and quadratic terms of the regression after normalization. The regression results are reported in Table E.6.[7]

A.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 SMM estimation requires that we choose moments that are sensitive to variations in the structural parameters. We now rationalize the moments that we choose to match.

Table A.1 reports the Jacobian matrix associated with the estimation of the baseline model. Each entry of the matrix reports the percentage change in a moment given a one percent increase in a parameter. This table gives us some indication on which data moments are the most informative in helping us identify each parameter:

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

(ii) Average profitability 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. This, in turn, leads to slightly higher average markups, profitability, and heterogeneity in markups across firms. The average and the standard deviation of leader relative quality increase accordingly.

(iii) An increase in either superstar innovation cost scale parameter χ or small firm innovation cost scale parameter ν reduces the aggregate R&D intensity and output growth rate. Since superstar innovation has a direct effect on the growth rate of the economy, the effect of χ

[6] The inverted-U result and the quantitative magnitude of the top point are robust to using the logarithm of SG&A (xsga).

[7] To further establish the robustness of our results, we conduct the hypothesis test proposed in Lind and Mehlum (2010) for regressions in Table E.6, 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 E.7 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.

TABLE A.1: IDENTIFICATION: JACOBIAN MATRIX

	λ	η	χ	ν	ζ	ϕ
Growth rate	0.778	-0.076	-0.263	-0.038	-0.511	2.309
R&D/GDP	-0.064	-0.137	-0.207	-0.074	-1.650	1.358
Advertising/GDP	-0.960	-0.331	0.034	-0.087	-1.599	0.328
Average markup	0.112	0.012	0.000	-0.002	-0.295	-0.011
Std. dev. markup	0.679	0.146	-0.007	0.014	-0.378	-0.115
Labor share	-0.023	0.006	-0.001	0.004	0.242	-0.005
Entry rate	0.000	0.000	0.000	0.000	0.000	0.000
Avg profitability	0.298	0.041	0.039	0.005	-0.906	-0.253
Avg leader rel. quality	0.519	0.146	-0.028	0.071	0.365	-0.174
Std. dev. leader rel. quality	0.290	0.182	-0.059	0.109	0.455	0.206
β(innovation, rel. sales)	0.288	0.023	-0.139	0.021	0.144	-0.229
Top point (innovation, rel. sales)	0.240	0.083	-0.004	-0.001	0.125	-0.124
β(advertising, rel. sales)	-0.739	-0.247	0.003	-0.008	-0.428	0.039
Top point (advertising, rel. sales)	0.410	0.179	0.016	-0.030	0.170	-0.082
	ϵ	χ_a	ϕ_a	γ	ψ	τ
Growth rate	0.267	0.006	0.060	-0.160	-0.067	-0.116
R&D/GDP	0.524	-0.008	0.099	-0.297	-0.130	-0.224
Advertising/GDP	0.652	-0.203	-0.607	-0.179	-0.152	-0.262
Average markup	0.017	-0.019	-0.051	-0.059	-0.003	-0.005
Std. dev. markup	-0.099	-0.062	-0.228	-0.279	0.025	0.043
Labor share	-0.030	0.011	0.021	0.026	0.007	0.011
Entry rate	0.000	0.000	0.000	0.000	0.000	1.000
Avg profitability	-0.022	-0.029	-0.043	-0.076	0.008	0.014
Avg leader rel. quality	-0.524	-0.015	-0.111	0.081	0.124	0.215
Std. dev. leader rel. quality	-0.687	-0.005	-0.130	-0.035	0.191	0.330
β(innovation, rel. sales)	-0.114	0.004	-0.115	-0.052	0.037	0.064
Top point (innovation, rel. sales)	0.006	-0.008	-0.064	0.042	-0.002	-0.004
β(advertising, rel. sales)	0.048	0.049	0.124	-0.509	-0.015	-0.025
Top point (advertising, rel. sales)	0.197	-0.037	-0.106	0.325	-0.052	-0.090

Notes: The table shows the Jacobian matrix associated with the estimation of the baseline model. Each entry of the matrix reports the percentage change in a moment given a one percent increase in a parameter.

on the output growth rate relative to its effect on R&D intensity is larger, whereas ν has a larger impact on R&D intensity. In addition, χ and ν have opposite implications for the level and dispersion of leader quality. Overall, larger χ tends to reduce the innovation of superstar firms, narrowing the quality gaps between the industry leader and other superstar firms. In comparison, 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.

(iv) The relative productivity of small firms ζ is identified very precisely by matching the average markup and the labor share. 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 lower labor share.

(v) As innovation policies in our estimated model are below unity, an increase in the R&D cost convexity parameters ϕ and ϵ reduces the innovation cost, which increases R&D intensity and the growth rate. These two parameters, however, have different implications for the inverted-U relationship between superstar innovation and market share. While ϵ strongly influences the linear coefficient of the innovation-market share regression, changes in ϕ affect both the linear coefficient and the location of the top point of the inverted-U relationship. The two parameters' effects on the standard deviation of leader relative quality are also opposite.

(vi) Intuitively, the three advertising-related data moments are most informative in helping 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 in the same direction, overall they are more sensitive to the changes in ϕ_a. The two parameters' effects on the linear term of the innovation-market share regression are also opposite.

(vii) 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.

(viii) Given all other parameter values, the value of ψ is set to normalize the measure of small firms m_t 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 (35) to get $\psi = \frac{\sum_{\Theta} v^e(\Theta)\mu(\Theta)}{2\tau}$.

B Additional Derivations and Proofs

B.1 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:

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

$$\tilde{p}_{cjt} = \tilde{y}_{cjt}^{-\frac{1}{\eta}} y_{jt}^{\frac{1}{\eta}-1} Y_t \tag{B.2}$$

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:

$$\frac{\partial p_{ijt}}{\partial y_{ijt}} y_{ijt} + p_{ijt} = \frac{w_t}{q_{ijt}} \tag{B.3a}$$

$$\frac{\partial p_{ijt}}{\partial \omega_{ijt}} y_{ijt} = \chi_a \phi_a \omega_{ijt}^{\phi_a - 1} Y_t \tag{B.3b}$$

Output Choices, Market Shares and Markups Let us work out the first condition, equation (B.3a). 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 (B.3a) 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:

$$\frac{1}{M_{ijt}} = \left(\frac{\eta-1}{\eta}\right) - \sigma_{ijt} \left[\left(\frac{\eta-\gamma}{\eta\gamma}\right) \left(\frac{y_{jt}}{\tilde{y}_{sjt}}\right)^{\frac{\gamma-1}{\gamma}} + \left(\frac{\gamma-1}{\gamma}\right) \right] \tag{B.4}$$

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:

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

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

(B.4), we have:

$$M_{ijt} = \left[\left(\frac{\eta-1}{\eta}\right) - \left(\frac{\gamma-1}{\gamma}\right)\sigma_{ijt} - \left(\frac{\eta-\gamma}{\eta\gamma}\right)\tilde{\sigma}_{ijt}\right]^{-1} \tag{B.5}$$

This is the expression for the markup written in the main text (equation (19)), where once again:

$$\sigma_{ijt} \equiv \frac{p_{ijt}y_{ijt}}{\tilde{p}_{cjt}\tilde{y}_{cjt} + \sum_{h=1}^{N_{jt}} p_{hjt}y_{hjt}} = \frac{\hat{\omega}_{ijt}\left(\sum_{h=1}^{N_{jt}}\hat{\omega}_{hjt}\left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{\left(\frac{\tilde{y}_{cjt}}{y_{ijt}}\right)^{\frac{\gamma-1}{\gamma}} + \left(\sum_{h=1}^{N_{jt}}\hat{\omega}_{hjt}\left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}} \tag{B.6}$$

$$\tilde{\sigma}_{ijt} \equiv \frac{p_{ijt}y_{ijt}}{\sum_{h=1}^{N_{jt}} p_{hjt}y_{hjt}} = \frac{\hat{\omega}_{ijt}}{\sum_{h=1}^{N_{jt}}\hat{\omega}_{hjt}\left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}} \tag{B.7}$$

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 (B.3b). 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:

$$\chi_a\phi_a\omega_{ijt}^{\phi_a-1} = \frac{\sigma_{ijt}}{\hat{\omega}_{ijt}}\left\{\frac{\partial\hat{\omega}_{ijt}}{\partial\omega_{ijt}} + \left(\sum_{h=1}^{N_{jt}}\frac{\partial\hat{\omega}_{hjt}}{\partial\omega_{ijt}}\left(\frac{y_{hjt}}{y_{ijt}}\right)^{\frac{\eta-1}{\eta}}\right)\left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right)\tilde{\sigma}_{ijt} - \left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right)\sigma_{ijt}\right]\right\} \tag{B.8}$$

Next, we compute the partial derivatives $\{\frac{\partial\hat{\omega}_{hjt}}{\partial\omega_{ijt}}\}_{h=1}^{N_{jt}}$. Using the definitions for $\{\hat{\omega}_{hjt}\}_{h=1}^{N_{jt}}$, we have:

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

Therefore, using equation (B.7):

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

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

Back into equation (B.8), we obtain:

$$\begin{aligned}\chi_a\phi_a\omega_{ijt}^{\phi_a-1} = \frac{\sigma_{ijt}}{\hat{\omega}_{ijt}}\Bigg\{&\frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}}\left(1-\frac{\hat{\omega}_{ijt}}{N_{jt}}\right)\\ &+\frac{\hat{\omega}_{ijt}}{1+\omega_{ijt}}\left(1-\frac{1}{N_{jt}}\frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right)\left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right)\tilde{\sigma}_{ijt}-\left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right)\sigma_{ijt}\right]\Bigg\}\\ = \frac{\sigma_{ijt}}{1+\omega_{ijt}}\Bigg\{&1-\frac{\hat{\omega}_{ijt}}{N_{jt}}+\left(1-\frac{1}{N_{jt}}\frac{\hat{\omega}_{ijt}}{\tilde{\sigma}_{ijt}}\right)\left[\left(\frac{\gamma-\eta}{\gamma(\eta-1)}\right)\tilde{\sigma}_{ijt}-\left(\frac{(\gamma-1)\eta}{\gamma(\eta-1)}\right)\sigma_{ijt}\right]\Bigg\} \qquad \text{(B.9)}\end{aligned}$$

This is the expression for the optimal advertising spending that we present in the main text (equation (24)).

B.2 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:

$$\begin{aligned}\ln(Y_t) &= \int_0^1 \frac{\gamma}{\gamma-1}\ln\left[\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}}+\left(\sum_{i=1}^{N_{jt}}\hat{\omega}_{ijt}y_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right]dj\\ &= \int_0^1 \frac{\gamma}{\gamma-1}\ln\left[\tilde{y}_{cjt}^{\frac{\gamma-1}{\gamma}}\left(1+\left(\sum_{i=1}^{N_{jt}}\hat{\omega}_{ijt}\left(\frac{y_{ijt}}{\tilde{y}_{cjt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right)\right]dj\\ &= \int_0^1 \left[\ln\left(\tilde{y}_{cjt}\right)+\frac{\gamma}{\gamma-1}\ln\left(1+\left(\sum_{i=1}^{N_{jt}}\hat{\omega}_{ijt}\left(\frac{y_{ijt}}{\tilde{y}_{cjt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right)\right]dj\\ &= \int_0^1 \left[\ln\left(\frac{q_{cjt}}{w_t^{rel}}\sigma_{cjt}\right)+\frac{\gamma}{\gamma-1}\ln\left(1+\left(\sum_{i=1}^{N_{jt}}\hat{\omega}_{ijt}\left(\frac{y_{ijt}}{\tilde{y}_{cjt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right)\right]dj\\ &= \int_0^1 \left[\ln\left(\frac{q_{cjt}}{w_t^{rel}}\right)+\frac{1}{\gamma-1}\ln\left(1+\left(\sum_{i=1}^{N_{jt}}\hat{\omega}_{ijt}\left(\frac{y_{ijt}}{\tilde{y}_{cjt}}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}\right)\right]dj\\ &= \int_0^1 \ln\left(\frac{q_{cjt}}{w_t^{rel}}\right)dj+\sum_{\Theta}f_t(\Theta)\mu_t(\Theta) \qquad \text{(B.10)}\end{aligned}$$

where $f_t(\Theta)$ is defined in equation (38), 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:

$$\begin{aligned} \ln(Y_{t+\Delta t}) - \ln(Y_t) &= -\ln(w_{t+\Delta t}^{rel}) + \ln(w_t^{rel}) + \ln(1+\lambda) \sum_{\Theta} p_{lit}(\Theta) \Delta t \mu_t(\Theta) \\ &\quad + \sum_{\Theta} \sum_{\Theta'} \left[f_t(\Theta') - f_t(\Theta) \right] p_t(\Theta, \Theta') \mu_t(\Theta) \Delta t + o(\Delta t) \end{aligned} \tag{B.11}$$

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

$$g_t = -g_{w^{rel},t} + \ln(1+\lambda) \sum_{\Theta} p_{lit}(\Theta) \mu_t(\Theta) + \sum_{\Theta} \sum_{\Theta'} \left[f_t(\Theta') - f_t(\Theta) \right] p_t(\Theta, \Theta') \mu_t(\Theta) \tag{B.12}$$

B.3 Proof of Proposition 1

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:

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

B.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 (B.10) to write:

$$Y_0 = \exp\left(\int_0^1 \ln(q_{cj0}) dj - \ln(w^{rel}) + \sum_{\Theta} f(\Theta) \mu(\Theta) \right) \tag{B.13}$$

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.[8] The initial level of consumption is then given by:

$$C_0 = Y_0 \cdot \frac{C_0}{Y_0} = Y_0 \left(1 + \int_0^1 \sum_{i=1}^{N_{j0}} \chi z_{ij0}^{\phi} \, dj + \int_0^1 \sum_{i=1}^{N_{j0}} \chi_a \omega_{ij0}^{\phi_a} \, dj + \int_0^1 m_0 \nu X_{kj0}^{\epsilon} \, dj + \psi e_0^2\right) \tag{B.14}$$

where we have used the aggregate resource constraint (equation (36)) at $t = 0$ on the right-hand side. The welfare of the representative household can be found by imposing BGP to equation (1):

$$W = \int_0^{+\infty} e^{-\rho t} \ln(C_t) \, dt = \frac{1}{\rho}\left(\ln(C_0) + \frac{g}{\rho}\right) \tag{B.15}$$

Using formulas (B.13) and (B.14), equation (40) 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:

$$W^B = \frac{1}{\rho}\left(\ln\left(C_0^A(1+\varsigma)\right) + \frac{g^A}{\rho}\right) \tag{B.16}$$

Solving for ς, we get:

$$\varsigma = \frac{C_0^B}{C_0^A} \exp\left(\frac{g^B - g^A}{\rho}\right) - 1 \tag{B.17}$$

[8] 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.

C Model Extensions

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

In this extension, we assume that, with probability $\delta \in [0,1]$, advertising turns out to be (unexpectedly) purely deceptive. The equilibrium conditions and allocations are the same as in the baseline model, and in particular, there is no change in consumer demand. 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.

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

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

where $f_t(\Theta)$ is defined in equation (38), and $h_t(\Theta)$ is the analogue of $f_t(\Theta)$ when advertising is deceptive, that is:

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

Notice that measured output (equation (C.1)) coincides with output in the baseline model (equation (B.10)) when $\delta = 0$.

The remaining objects necessary to calculate welfare are computed as in the baseline model, as explained in Section B.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 (B.14), that is:

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

C.2 Non-Combative Advertising

Under the specification in equation (41), the demand schedules faced by every firm are unchanged, so that equations (B.1)-(B.2) still hold. Likewise, the optimal firm-level markup is still given by equation (B.5). The advertising choice, however, is slightly different. While equation (B.8) 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 (41), we have:

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

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

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

Therefore:

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

Plugging back into equation (B.8), we obtain:

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

Note that setting $\Lambda = 0$, we return to the optimality condition of the baseline model (equation (B.9)). For $\Lambda = 1$, we obtain:

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

C.3 Bertrand model

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:

$$y_{ijt} + \left(p_{ijt} - \frac{w_t}{q_{ijt}}\right) \frac{\partial y_{ijt}}{\partial p_{ijt}} = 0 \tag{C.2a}$$

$$\left(p_{ijt} - \frac{w_t}{q_{ijt}}\right) \frac{\partial y_{ijt}}{\partial \omega_{ijt}} = \chi_a \phi_a \omega_{ijt}^{\phi_a - 1} Y_t \tag{C.2b}$$

Price Choice Let us first work out condition (C.2a). 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:

$$\frac{\partial y_{ijt}}{\partial p_{ijt}} = -\eta \frac{y_{ijt}}{p_{ijt}} + \underbrace{\hat{\omega}_{ijt}^{\eta} p_{ijt}^{-\eta} \tilde{p}_{sjt}^{\eta-\gamma} p_{jt}^{\gamma-1}}_{=y_{ijt}/Y_t} \left[(\eta-\gamma) \left(\frac{\tilde{p}_{sjt}}{p_{jt}}\right)^{\gamma-1} y_{ijt} + (\gamma-1) y_{ijt} \right] \tag{C.3}$$

Using $\frac{\partial p_{ijt}}{\partial y_{ijt}} y_{ijt} + p_{ijt} = \frac{w_t}{q_{ijt}}$ by equation (C.2a) 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 (C.3) that:

$$\mathcal{E}_{ijt} = -\frac{p_{ijt}}{y_{ijt}} \frac{\partial y_{ijt}}{\partial p_{ijt}} = \eta - \sigma_{ijt} \left[(\eta-\gamma) \left(\frac{\tilde{p}_{sjt}}{p_{jt}}\right)^{\gamma-1} + (\gamma-1) \right] \tag{C.4}$$

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:

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

Advertising Choice Next, we move to the optimality condition for advertising effort, equation (C.2b). First, we have:

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

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:

$$\begin{aligned}
\chi_a \phi_a \omega_{ijt}^{\phi_a-1} = \frac{\sigma_{ijt}(1-M_{ijt}^{-1})}{\hat{\omega}_{ijt}^{\eta}} \Bigg\{ & \eta \hat{\omega}_{ijt}^{\eta-1} \frac{\partial \hat{\omega}_{ijt}}{\partial \omega_{ijt}} \\
& + \left(\sum_{h=1}^{N_{jt}} \hat{\omega}_{hjt}^{\eta-1} \frac{\partial \hat{\omega}_{hjt}}{\partial \omega_{ijt}} \left(\frac{p_{hjt}}{p_{ijt}}\right)^{1-\eta} \right) \left[\left(\frac{\eta(\gamma-\eta)}{\eta-1}\right) \tilde{\sigma}_{ijt} - \left(\frac{\eta(\gamma-1)}{\eta-1}\right) \sigma_{ijt} \right] \Bigg\}
\end{aligned} \tag{C.5}$$

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:

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

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

Substituting this back into equation (C.5), we obtain:

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

D 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.

D.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:

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

subject to

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\frac{\dot{Q}_t}{Q_t} = \ln(1+\lambda) \sum_{\Theta} p_{lit}(\Theta) \mu_t(\Theta) \tag{D.2q}$$

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

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

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

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

D.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:

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

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

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

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

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

where $\vartheta_t > 0$ is the Lagrange multiplier associated with the labor feasibility constraint (D.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:

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

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

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,

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

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

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

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

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 (D.9), some algebra shows:

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

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

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

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

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

$$\left(\frac{\eta}{\eta-1}\right)\frac{\left(\sum_{k=1}^{N_{jt}}\hat{\omega}_{kjt}y_{kjt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\gamma-\eta}{\gamma(\eta-1)}}}{y_{cjt}^{\frac{\gamma-1}{\gamma}}+\left(\sum_{k=1}^{N_{jt}}\hat{\omega}_{kjt}y_{kjt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta(\gamma-1)}{\gamma(\eta-1)}}}\frac{N_{jt}}{\left(\sum_{k=1}^{N_{jt}}(1+\omega_{kjt})\right)^2}\left[y_{ijt}^{\frac{\eta-1}{\eta}}\sum_{h\neq i}(1+\omega_{hjt})-\sum_{h\neq i}(1+\omega_{hjt})y_{hjt}^{\frac{\eta-1}{\eta}}\right] = \frac{\chi_a\phi_a\omega_{ijt}^{\phi_a-1}}{1-A_t} \tag{D.15}$$

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

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

D.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:

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

In equation (D.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 (20)) can be written in terms of markups as follows:

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

$$\sigma_{cjt} = \frac{1}{1 + \left(\sum_{k=1}^{N_{jt}} \hat{\omega}_{kjt}^{\eta} \left(\frac{q_{kjt}}{q_{cjt}} \right)^{\eta-1} \left(\frac{M_{cjt}}{M_{kjt}} \right)^{\eta-1} \right)^{\frac{\gamma-1}{\eta-1}}}, \quad \forall j \in [0,1] \tag{D.19}$$

Comparing allocation (D.17) with the social planner's (equations (D.12)-(D.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).

E Additional Figures and Tables

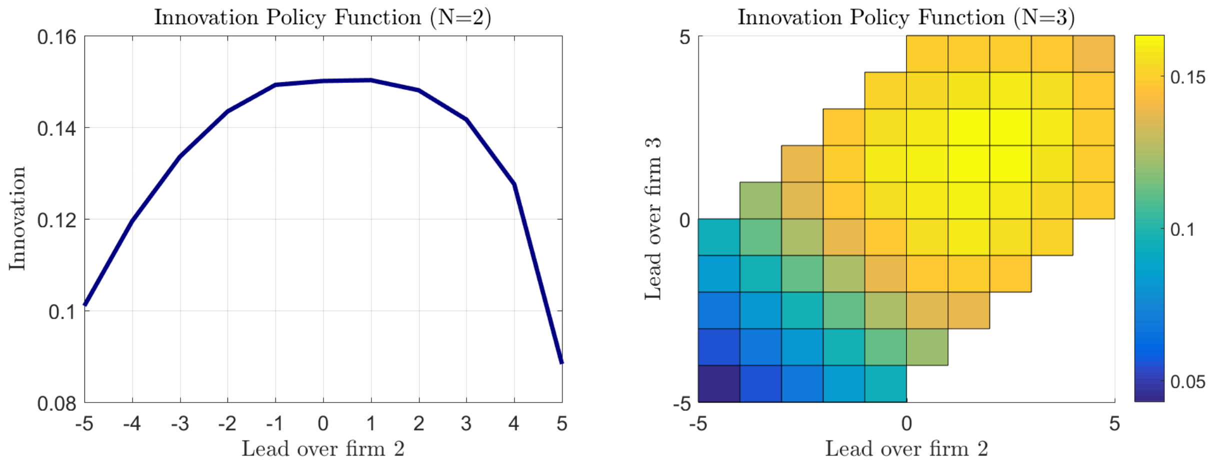

FIGURE E.1: INNOVATION POLICY FUNCTION

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.

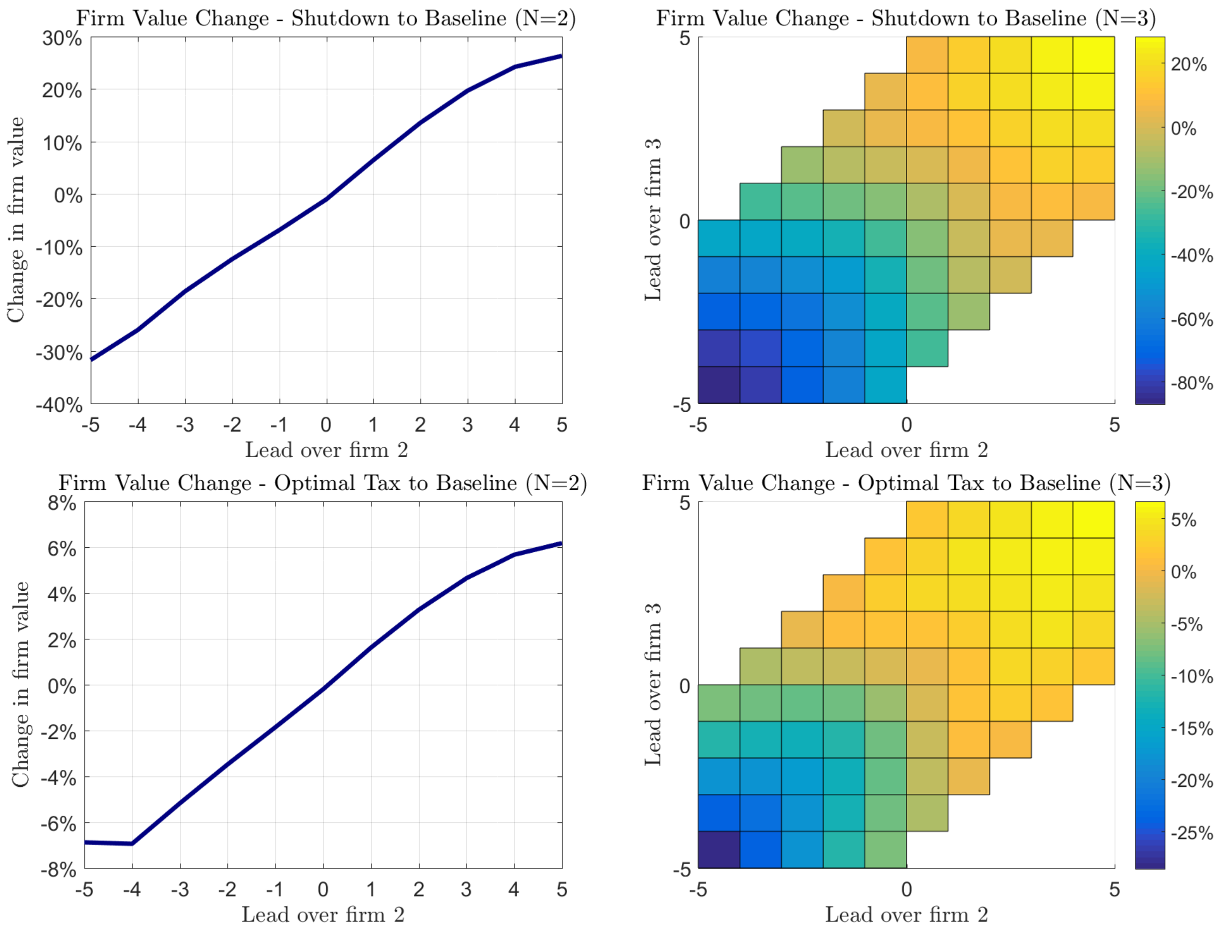

FIGURE E.2: CHANGE IN FIRM VALUE FROM SHUTDOWN AND OPTIMAL TAX

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.

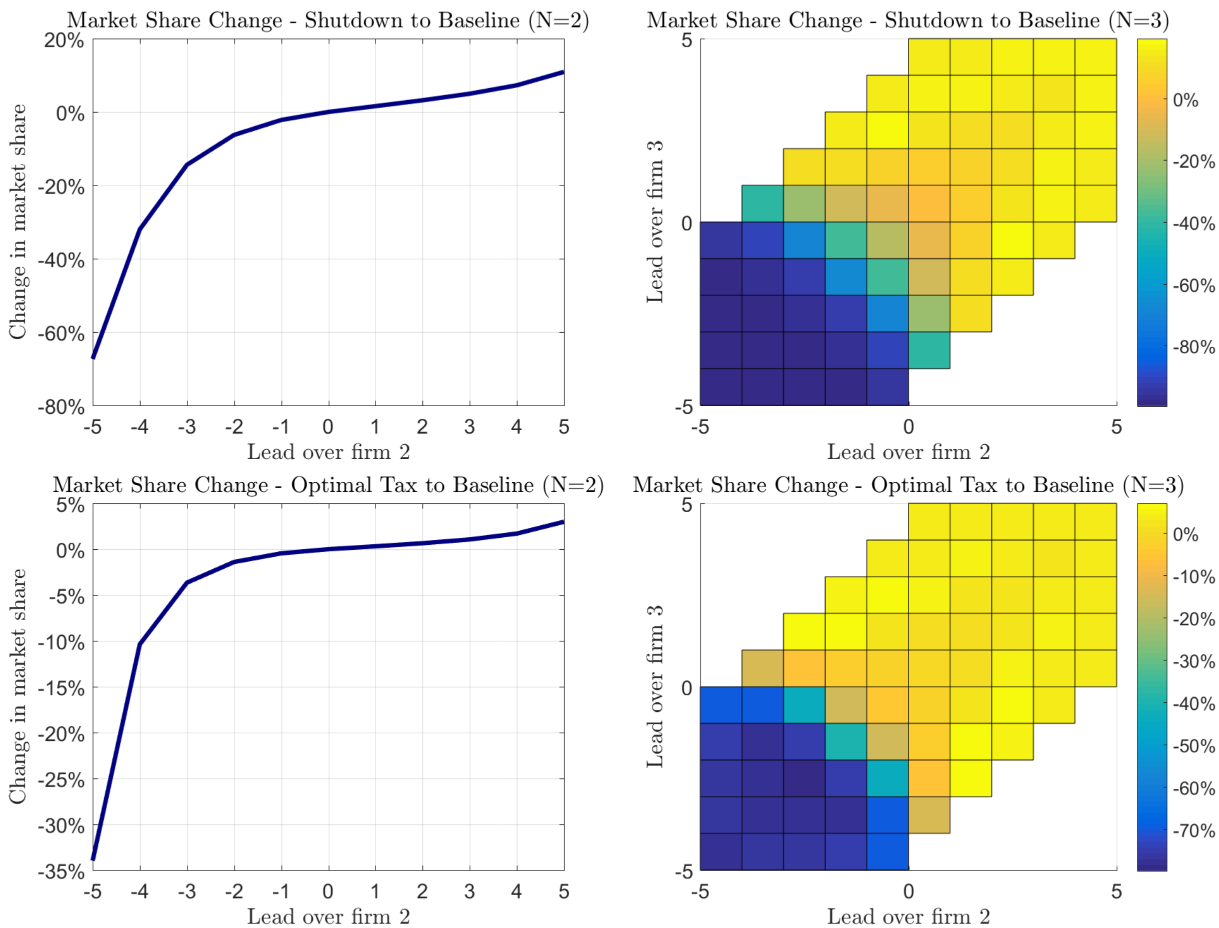

Figure E.3: Change in Market Shares from Shutdown and Optimal Tax

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.

Table E.1: 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 E.2: EXTENDED MODEL PARAMETERS AND TARGET MOMENTS (NON-COMBATIVE ADVERTISING $\Lambda = 1$)

A. Parameter estimates

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
β(innovation, relative sales)	0.629	1.043
Top point (innovation, relative sales)	0.505	0.475
β(advertising, relative sales)	6.260	7.544
Top point (advertising, relative sales)	0.533	0.563

Notes: The estimation is done with the Simulated Method of Moments. Panel A reports the estimated parameters. Panel B reports the simulated and empirical moments.

Table E.3: The Impact of Advertising Shutdown and Optimal Advertising Tax with Non-Combative Advertising ($\Lambda = 1$)

	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 E.4: EXTENDED MODEL PARAMETERS AND TARGET MOMENTS (BERTRAND COMPETITION)

A. Parameter estimates

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
β(innovation, relative sales)	0.629	0.821
Top point (innovation, relative sales)	0.505	0.433
β(advertising, relative sales)	6.260	8.581
Top point (advertising, relative sales)	0.533	0.499

Notes: The estimation is done with the Simulated Method of Moments. Panel A reports the estimated parameters. Panel B reports the simulated and empirical moments.

Table E.5: The Impact of Advertising Shutdown and Optimal Advertising Tax (Bertrand Competition)

	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 E.6: 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 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 E.7: 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 testing results.

Table E.8: 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.