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

Laurent Cavenaile	Murat Alp Celik	Pau Roldan-Blanco	Xu Tian
University of Toronto	University of Toronto	Bank of Spain	University of Georgia

April 29, 2022

Abstract

Firms use both innovation and advertising to increase their profits, markups, and market shares. While they serve the same purpose from the firms' perspective, their broader implications vary substantially. In this paper, we study the interaction between these two intangible inputs and analyze the implications for competition, industry dynamics, economic growth, and social welfare. To this end, 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. We estimate the model to fit the non-linear relationship between innovation, advertising, and competition observed in the data. We find that advertising has significant macroeconomic effects: it improves static allocative efficiency through reducing misallocation, but it also depresses economic growth through a substitution effect with R&D. On the net, advertising is found to be welfare-improving. It is responsible for one quarter of the observed average net markup and its dispersion. We next study the optimal linear taxation/subsidization of advertising. We find that the optimal advertising tax is quite high. Such taxation could simultaneously increase dynamic efficiency, contain excessive spending on advertising due to inefficient "rat race", and raise revenue while still maintaining most of the benefits of advertising via improving efficiency in resource allocation.

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

JEL Classification: E20, L10, M30, O30, 040.

*Contact Information: Laurent Cavenaile, 105 St. George Street, Rotman School of Management, University of Toronto, Toronto, Ontario, M5S 3E6. Email: laurent.cavenaile@utoronto.ca; Murat Alp Celik, 229 Max Gluskin House, 150 St. George Street, University of Toronto, Toronto, Ontario, M5S 3G7. Email: murat.celik@utoronto.ca; Pau Roldan-Blanco, Calle Alcalá 48, 28014, Madrid. Email: pau.roldan@bde.es; Xu Tian, 620 South Lumpkin Street, Department of Finance, Terry College of Business, University of Georgia, Athens, GA 30602. Email: xu.tian@uga.edu.

We would like to thank Michael Klein, Pedro Monteiro, Sergio Ocampo, Miquel Oliver i Vert, Jesse Perla, Nicolas Sahuguet, Kenneth Simons, Peifan Wu, 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 Autonoma de Barcelona, Rensselaer Polytechnic Institute, and Econometric Society Winter Meeting 2021 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).

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. However, advertising is another activity through which firms can achieve the same desired outcomes, albeit without contributing to long-run productivity growth. On the contrary, firms might spend exorbitant amounts in advertising in response to the advertising of their competitors, leading to an inefficient "rat race" equilibrium with excessive spending. Since both activities serve the same purpose, 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 into both innovation and advertising. Since 1980, R&D accounts for 2.44% of the U.S. GDP, 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 simply marketing them ("style"). For instance, Procter & Gamble Company has 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 in advertising is socially efficient, and (2) if the 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, 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. 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 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 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

[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.

and gain access to the group of large firms. This rich yet tractable setting allows us to tackle relevant policy questions. On the one hand, 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. On the other hand, the model allows us to 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 the model, R&D and advertising are modeled as intangible expenditures which can shift demand toward the firm's product. R&D is modeled following the tradition of the Schumpeterian creative-destruction literature. We model advertising to be 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. 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 firm size can magnify productivity differences and have quantitatively significant implications for within-industry markup dispersion and, as a consequence, allocative efficiency. Moreover, because in equilibrium firms are heterogeneous in their use of advertising, there is a dynamic interplay between advertising and R&D decisions, which at the aggregate level has an impact on the rate of economic growth and social welfare.

To study these questions quantitatively, we estimate the model by 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. 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 exert 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, consistent with the empirical findings in Cavenaile and Roldan-Blanco (2021). 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 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; as well as amplify 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, in the relative industry output of large firms, in the consumption share of GDP, and in the rate of economic growth. 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. All in all, 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, in the last part of the paper, we ask how the design of policy instruments should view these welfare results. Although we conclude that shutting down advertising would reduce dynamic efficiency, 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 the firms from investing resources in advertising, resulting in both direct gains in the consumption-to-output ratio, and indirect gains from improved incentives for innovation and growth. However, the taxes do not cause as large a drop in static allocative efficiency as a complete shutdown would: while the overall spending on advertising declines, more productive superstars still continue to spend more on advertising than less productive ones. Therefore, the positive effects of advertising due to the more efficient reallocation of resources 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 a perfect candidate for taxation to raise revenues while simultaneously increasing dynamic efficiency.

Literature Review Our paper is primarily related to a literature that studies the implications of intangible investments, primarily in the form of advertising and customer capital, for firm, industry and macroeconomic dynamics (e.g. Gourio and Rudanko (2014), Molinari and Turino (2017), Perla (2019), Dinlersoz and Yorukoglu (2012), and Greenwood, Ma, and Yorukoglu (2020)). The literature has investigated, for instance, how intangibles may be behind several trends related to business dynamism, market concentration and markups (e.g. Aghion, Bergeaud, Boppart, Klenow, and Li (2019), DeRidder (2020) and Weiss (2019)), or how they may affect markup cyclicality (Roldan-Blanco and Gilbukh (2021)), firm's market value and risk (Belo, Lin, and Vitorino (2014)), the transmission channels of monetary policy (Morlacco and Zeke (2020)), and the behavior of exporters and international prices (Drozd and Nosal (2012), Fitzgerald, Haller, and Yedid-Levi (2017)). Our model is most closely related to endogenous growth models with advertising such as Rachel (2021), Cavenaile and Roldan-Blanco (2021), and Klein and Sener (2021). Rachel (2021) studies how the provision of free leisure-enhancing technologies that firms can use to build their brand equity (e.g. through advertising) can explain the observed decline in hours workers and affect negatively innovation and TFP growth.[2] In a model with monopolistic competition,

[2] Greenwood, Ma, and Yorukoglu (2020) 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.

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 Sener (2021) study how both informative and combative advertising affect the speed of diffusion of innovation. Studying policy implications, they find that R&D subsidies increase innovation rates but decreases advertising and diffusion leading to an ambiguous effect on growth and welfare. In the present paper, we explore how the interaction between R&D and advertising further affects market concentration, markups, productivity growth, and welfare. As in Cavenaile and Roldan-Blanco (2021) and Klein and Sener (2021), we find that advertising and R&D are substitutes, even though the three theoretical frameworks are different.

Our paper also contributes to the growing literature studying heterogeneous-firm economies with markup distortions (e.g. Edmond, Midrigan, and Xu (2018), Burstein, Carvalho, and Grassi (2020), Baqaee and Farhi (2020)). Within this literature, we are most closely related to models of endogenous growth with variable markups, e.g. Akcigit and Ates (2019a,b), Peters and Walsh (2019), Liu, Mian, and Sufi (2019), Aghion, Bergeaud, Boppart, Klenow, and Li (2019), Peters (2020) and, particularly, Cavenaile, Celik, and Tian (2020). In our model, as in models of oligopolistic competition in the spirit of Atkeson and Burstein (2008), firm-level markups are increasing in the market share of firms. However, as in Cavenaile, Celik, and Tian (2020), market structure is endogenous. In addition, there is an interplay between market shares and firms' innovation decisions. To this framework, we add advertising decisions at the firm level, which in equilibrium directly affect market shares and markups as well as R&D decisions.

More generally, our paper contributes to a long tradition of modeling advertising in economics (e.g. Dorfman and Steiner (1954), Butters (1977), Becker and Murphy (1993), Benhabib and Bisin (2002)).[3] In the literature, advertising mostly acts as a demand shifter by affecting demand preferences. Similarly, 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

[3] Bagwell (2007) provides a comprehensive survey of the advertising literature in economics.

relationships emerge endogenously, and help explain the macroeconomic implications of R&D and advertising on growth and welfare.

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 of the model and 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). Section 5 offers some concluding remarks.

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 a 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 homogenous 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. Total production by superstars of industry j at time t is given by:

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

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 effective 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 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, but so that if all superstars were to choose the same ω level, then all varieties would have the same baseline quality. This baseline quality coincides with the quality of the product of the fringe, which is normalized to one.

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

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

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

Advertising Each superstar firm can spend resources on advertising its product to affect the quality shifter $\hat{\omega}_{ijt}$. In order to achieve a perceived quality 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 rate 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, the number 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 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. 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 standard 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:

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

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

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 homogenous 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 allocation of expenditures among superstar firms of the same industry is determined by perceived product qualities $\hat{\omega}_{ijt}$ as well as the price of each firm's product relative to other superstars in its industry, $\frac{p_{ijt}}{\tilde{p}_{sjt}}$, with price-elasticity η. In particular, 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. 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 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 choose 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 $y_{ijt} = q_{ijt}l_{ijt}$. We assume that superstar firms within the same industry compete *à la* Cournot, so they internalize how their output choices affect the aggregate output within their industry. In equilibrium, each superstar firm i sets a price 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.[4] Equation (19) shows that a firm's markup is increasing in both of these market share measures. Importantly, we

[4]Indeed, the numerators in Equation (20) are firm sales while, by equation (17), the denominator of σ (respectively, $\tilde{\sigma}$) equals the industry's expenditures including (respectively, excluding) the fringe.

can write market shares in terms of relative outputs and productivities, as follows:[5]

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

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 its relative productivity and relative taste shifter, and decreasing in its 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} in equation (18), 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

[5]In fact, a simple mapping between σ and $\tilde{\sigma}$ is $\tilde{\sigma}_{ijt}\tilde{y}_{sjt}^{\frac{\gamma-1}{\gamma}} = \sigma_{ijt}y_{jt}^{\frac{\gamma-1}{\gamma}}$.

be solved, for each industry j, as a function of the set of relative productivities between firms, $\{n_{ijt}^k\}$, and the total number of superstars in the industry, N_j. Though 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 $\pi_{ijt}^{adv} \equiv \pi_{ijt} - \chi_a \omega_{ijt}^{\phi_a} Y_t$, 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:

$$l_{ijt} = \frac{\sigma_{ijt}}{w_t^{rel}} M_{ijt}^{-1} \quad \text{and} \quad l_{cjt} = \frac{\sigma_{cjt}}{w_t^{rel}} \tag{25}$$

for each superstar firm i and the fringe, respectively, where $w_t^{rel} \equiv \frac{w_t}{Y_t}$ 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:[6]

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

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 of 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, $\mathbf{n}_{ijt} = \{n_{ijt}^k\}_{k \neq i}$, 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:

$$\begin{aligned} rV(\mathbf{n}_i, N) = \max_{z_i} \Bigg\{ & \pi^{adv}(\mathbf{n}_i, N) - \chi z_i^{\phi} Y \\ & + z_i \left[V\left(\mathbf{n}_i \backslash \{n_i^k = \bar{n}\} + \mathbf{1}, N - |\{n_i^k = \bar{n}\}|\right) - V(\mathbf{n}_i, N) \right] + \sum_{\{k: n_i^k = -\bar{n}\}} z_{kj} \Big(0 - V(\mathbf{n}_i, N) \Big) \end{aligned}$$

[6]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 harmonic average of firm-level markups.

$$+\sum_{\{k:n_i^k>-\bar{n}\}} z_{kj}\left[V\Big(\mathbf{n}_i\backslash\{n_i^k\}\cup\{n_i^k-1\}\backslash\{n_i^l=\bar{n}+n_i^k\},N-|\{n_i^l=\bar{n}+n_i^k\}|\Big)-V(\mathbf{n}_i,N)\right]$$

$$+X_j\left[V(\mathbf{n}_i\cup\{\min\{\bar{n},\bar{n}+\min(\mathbf{n}_i)\}\},\min(N+1,\bar{N}))-V(\mathbf{n}_i,N)\right]\Bigg\}+\dot{V}(\mathbf{n}_i,N) \tag{27}$$

In this Hamilton-Jacobi-Bellman equation, the first line is the flow profit from sales net of labor and advertising costs, minus the costs from R&D. The firm term on the second line is the gain from a successful innovation, which increases the lead of the firm by one 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 firms 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 entry 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 value due to firm growth.

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:

$$\rho v(\mathbf{n}_i,N)=\max_{z_i}\Bigg\{\frac{\pi^{adv}(\mathbf{n}_i,N)}{Y}-\chi z_i^{\phi}+z_i\left[v\Big(\mathbf{n}_i\backslash\{n_i^k=\bar{n}\}+\mathbf{1},N-|\{n_i^k=\bar{n}\}|\Big)-v(\mathbf{n}_i,N)\right]$$

$$+\sum_{\{k:n_i^k\neq-\bar{n}\}} z_{kj}\left[v\Big(\mathbf{n}_i\backslash\{n_i^k\}\cup\{n_i^k-1\}\backslash\{n_i^l=\bar{n}+n_i^k\},N-|\{n_i^l=\bar{n}+n_i^k\}|\Big)-v(\mathbf{n}_i,N)\right]$$

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

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 $p_{li}(\Theta)$ as the arrival rate of a leader innovation and $p(\Theta, \Theta')$ 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:

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

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.[7] 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:

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

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.[8] The value of being an entrepreneur (S) is:

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

[7]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.

[8]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.2). W is thus equal to $\frac{1 - \frac{1}{\epsilon}}{\rho + \tau} (\nu \epsilon)^{\frac{1}{1 - \epsilon}} \int_0^1 V(\{\tilde{\mathbf{n}}_j - \bar{n}\} \cup \{-\bar{n}\}, N_j + 1)^{\frac{\epsilon}{\epsilon - 1}} \, dj$.

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 An equilibrium 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, 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 profit.

(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, incumbents' R&D, incumbents' 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:[9]

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

where $g_{w^{rel},t}$ 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:

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

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:

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

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.

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 markups at different levels of aggregation.

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.[10] Matching both of these margins helps carefully discipline the

[9] See Appendix B.1 for the full derivation.

[10] See Cavenaile, Celik, and Tian (2020) for the documentation of both regularities, and their robustness across different specifications.

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.1 all the 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's moments with respect to each estimated parameter, see Appendix A.2.

Table 1 presents the results of our SMM estimation exercise. Panel A reports the parameter values, and Panel B reports the results in terms of moment matching. 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 help 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.

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	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.

3.2 Optimal Advertising and Innovation Policies

Using the calibrated 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). 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.

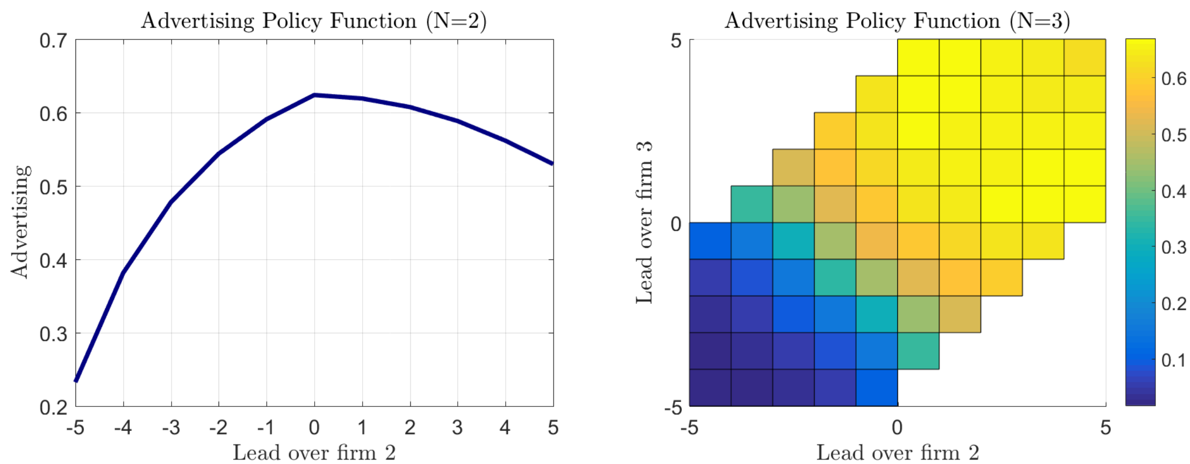

FIGURE 1: ADVERTISING POLICY FUNCTION

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 exhibit a similar pattern, with advertising incentives increasing in technological lead, and declining (though only slightly) when the firm is far ahead of both of its competitors.[11] Figure D.1 in the Appendix shows the corresponding policy functions for innovation, exhibiting a similar feature: firms innovate the most when they are close to being neck-to-neck, and innovation incentives decrease the higher the technological with their competitors.

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

[11] Since we assume that there exists 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.

sure that the model can reproduce the empirically observed relationship between innovation, advertising, and competition.

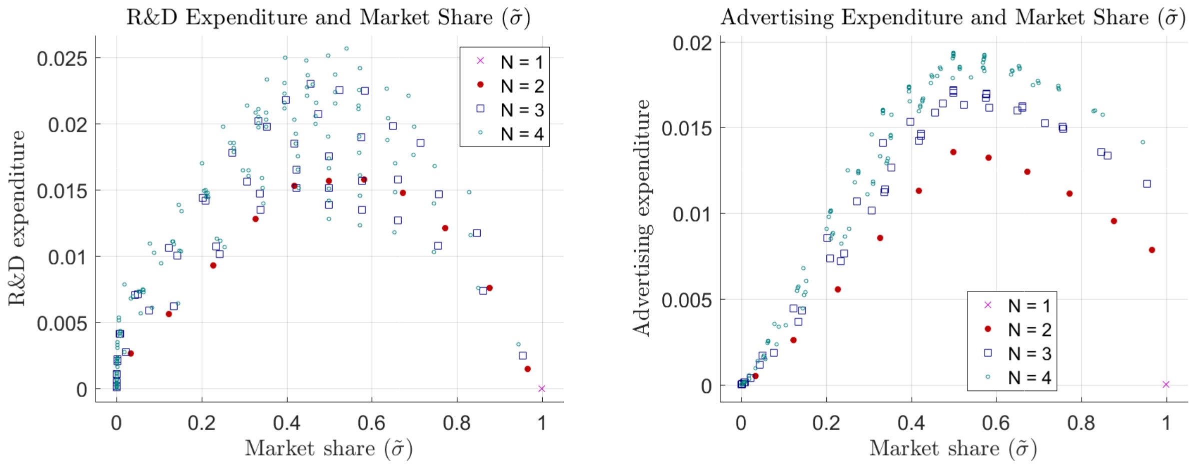

FIGURE 2: R&D EXPENSES, ADVERTISING, AND FIRM MARKET SHARES

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 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 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.

Figure 3 shows the model-implied cross-industry relationship between advertising and competition. In particular, we plot the total industry advertising expenditure as a function of the industry's Herfindahl-Hirschman Index (HHI). Each circle is an industry, 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 invariant distribution $\mu(\Theta)$. We observe that industries with more superstar firms have, on average, a higher overall investment in advertising. Interestingly, more concentrated industries (as measured by the HHI) are characterized, on average, by higher advertising expenditures.

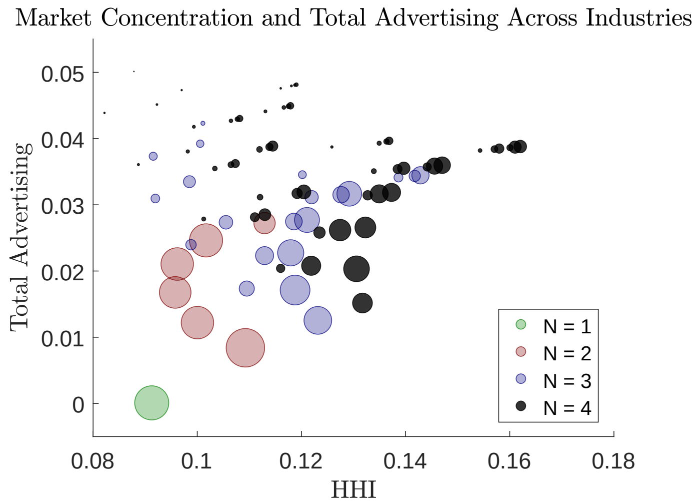

FIGURE 3: ADVERTISING AND MARKET CONCENTRATION ACROSS INDUSTRIES

3.4 The Relationship Between Markups, Advertising, and Innovation

In Section 2.2, we found that markups are increasing in market shares, and that advertising decisions are driven by endogenous changes in market power through both markups and market shares. To further understand the relationship between competition, innovation, and advertising for the estimated set of parameters, Table 2 reports a matrix of correlation coefficients at the firm level for markups, R&D expenditures, advertising expenditures, the sum of the two (which we label, following Compustat nomenclature, as Selling, General and Administrative expenses, or SG&A), and profitability.[12]

TABLE 2: 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

In the table, we observe that all of these variables are positively correlated with one another in our estimated model. First, R&D and advertising are strongly positively correlated at the firm level: those firms that invest the most in R&D also spend relatively more in advertising. Second, both R&D and advertising are positively correlated with markups,

[12]Profitability is defined as static profits minus R&D costs over sales.

though advertising more so than R&D. This is in line with empirical evidence presented in De Loecker, Eeckhout, and Unger (2020), who find a positive firm-level relation between markups and both R&D and advertising expenditures for publicly-traded firms. This suggests that both in the model as in the data, firms use both advertising and innovation to shift demand away from their competitors and into their products, allowing them to increase their profits and charge higher markups. Indeed, there is also a positive correlation between R&D, advertising, and profitability, indicating that by increasing their R&D and advertising expenditures, firms wield greater market power within their industries.

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. Finally, we study the policy implications of our model by considering the linear taxation/subsidization of advertising.

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.[13] 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 3 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 product. On the one hand, advertising allows firms to magnify

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

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 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 suggests that the second effect dominates, and that R&D and advertising are substitutes at the aggregate level in general equilibrium, as innovation by superstars increases in response to shutting down advertising. 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, in which 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 to become superstars themselves. Overall, shutting down advertising raises economic growth by 3.26% of its value. This result is in line with the results in Cavenaile and Roldan-Blanco (2021). 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 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 largest superstars.

In line with the correlations presented in Table 2, 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 markup decreases from 1.34 to 1.25. 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 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.

TABLE 3: 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 (after-tax)	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%

The decrease in the average markup is accompanied with a decline in the profitability of superstar firms (-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 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. We analyze this effect in this section.

The analysis in Appendix C.3 shows how advertising affects static allocative efficiency. For a given industry state, it affects industry output by (i) influencing the individual output (y_{ij}) of firms with different productivities, (ii) modifying the perceived quality ($\hat{\omega}_{ij}$) of different varieties, and (iii) changing the relative wage.

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 efficient superstars (high q_{ij}). This means that the economy with advertising uses resources 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 (or cheaper to produce) varieties. Combined together, the reallocation of resources towards more efficient firms, and the relative amplification of the perceived quality of the same, work against the effects of higher markups; implying that advertising helps improve static allocative efficiency on the net.

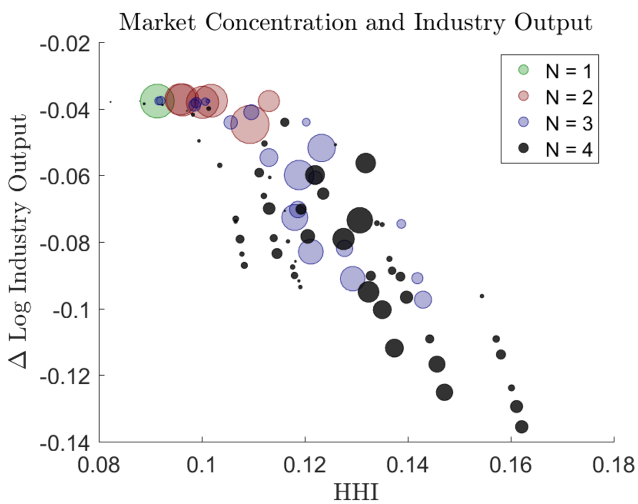

FIGURE 4: CHANGE IN INDUSTRY OUTPUT BY MARKET CONCENTRATION

Focusing on how shutting down advertising affects allocative efficiency across different industry states also reveals interesting patterns. Figure 4 depicts the difference in industry output between our baseline model and our counterfactual economy as a function of industry concentration measured by the HHI (taken from the baseline economy). It shows that the decrease in industry output is larger for industries that are more concentrated. Advertising allows larger firms to shift demand towards their products and away from those smaller and less productive. This is further clarified in the left panel of Figure 5, which displays the change in industry output as a function of productivity dispersion in the industry. As seen in the figure, this reallocation of market shares, improving static allocative efficiency, is stronger in industries where the dispersion in terms of productivity is larger. The right panel of Figure 5 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, but despite its positive effects, the reallocation and perceived quality amplification channels dominate, and we observe an overall decline in industry output across the board for all industry states.

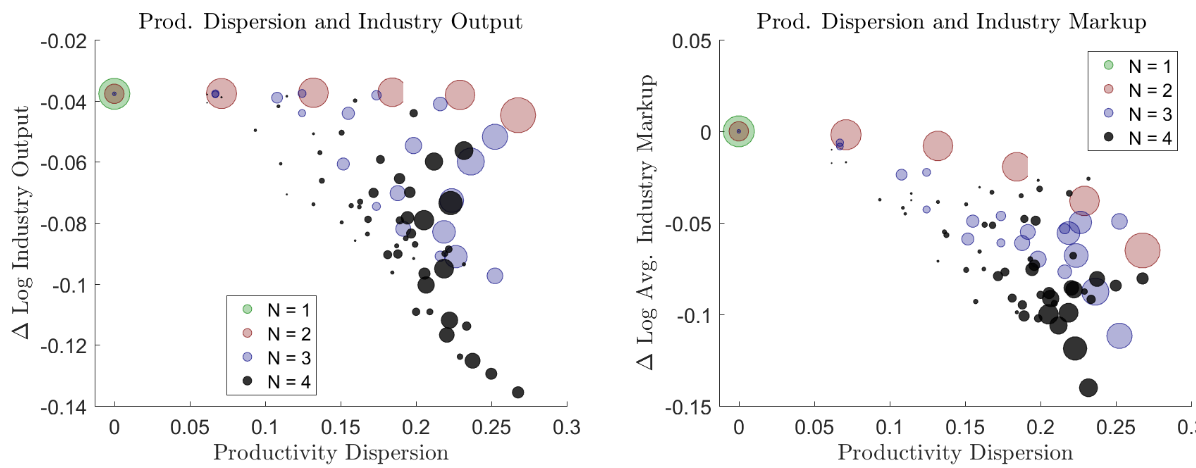

FIGURE 5: CHANGE IN INDUSTRY OUTPUT AND MARKUP BY PRODUCTIVITY DISPERSION

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. 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 firm 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 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.3):

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

The first term in the square brackets corresponds to the change in the relative productivity of the fringe across the two economies, the second term reflects the change in the relative wage, and the third term relates to changes in the relative industry output of superstar firms. These three terms collectively represent the change in welfare due to the change in the initial output level, Y_0. The fourth term captures changes in the consumption share of GDP. The last term in the equation captures how the differential in the growth rates between the two economies is translated to changes in welfare.

TABLE 4: ADVERTISING SHUTDOWN: SHORT-RUN VS. LONG-RUN EFFECTS ON EFFICIENCY

	Static		Static+New Distribution		Dynamic	
	ΔW	CEWC	ΔW	CEWC	ΔW	CEWC
Competitive fringe productivity	0.000	0.00%	0.000	0.00%	0.000	0.00%
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%

Table 4 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.[14] 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 decrease in the relative wage and in 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 4 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 4 show the full results including the dynamic effects due to changes in R&D investment. Shutting down advertising raises the consumption-to-output ratio as a result of changes in total R&D and advertising expenses. In addition, the growth rate of the economy increases. Overall, these dynamic effects further offset some of the static welfare losses. All in all, static losses are larger than dynamic gains, resulting in a total welfare loss of 0.86% in consumption equivalent terms when advertising is shut down. Therefore, we conclude that advertising, despite its various negative effects, helps rather than hurts efficiency, albeit by a small 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 5 reports the results of our policy experiment for different values of taxes and subsidies.[15] In line with the results of our shutdown experiment, higher taxes (subsidies) on advertising are associated with a reduction (increase) in advertising expenditures and

[14] 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.8) of Appendix B.3.

[15] 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.

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.

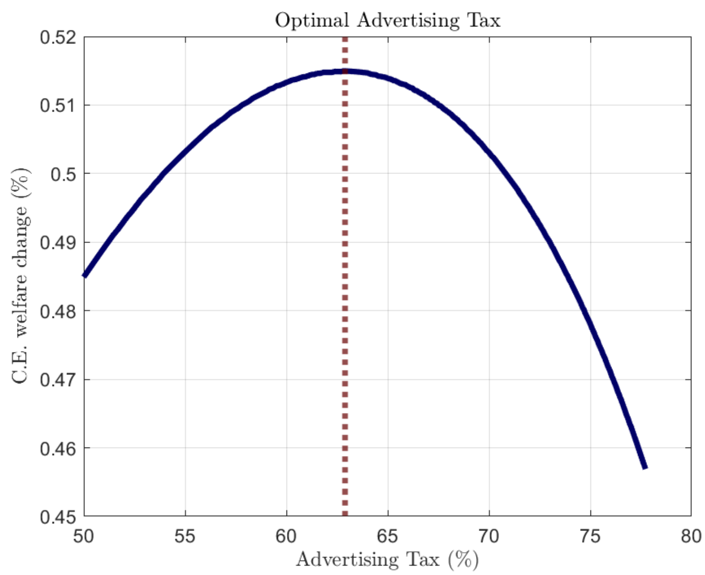

FIGURE 6: THE DYNAMIC WELFARE IMPACT OF ADVERTISING TAXES

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 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 quite high at 62.9%? The answer lies in understanding how taxation differs from a complete shutdown. Higher taxes on advertising expenses discourage the 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, the taxes do not cause as large a drop in static allocative efficiency as a complete shutdown would: while the overall spending on advertising declines, more productive superstars still continue to spend more on advertising than less productive ones. Therefore, the positive effects of advertising due to the more efficient reallocation of resources 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 a perfect candidate for taxation.

In most advanced economies including the United States, advertising expenses are not taxed. Our quantitative analysis demonstrates that advertising is a useful activity insofar that 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.[16] 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. 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.

[16]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.

Table 5: 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%	

5 Conclusion

Firms routinely make intensive use of their innovation and advertising expenditures in order to alter 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 make up for a large share of GDP in the United States. Yet, the interaction between them and their implications for economic growth and social welfare remain understudied in the economics 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 share of the industry's total output), 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 gain access to 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. Specifically, shutting down advertising leads to a decrease in allocative efficiency, because advertising reallocates market shares toward larger and more productive firms. However, as in the estimated model, advertising and innovation are substitutes, shutting down advertising frees up resources for firms to use in their innovation expenditures, 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 helps static allocative efficiency through reallocating resources towards more efficient firms, and its shutdown therefore reduces static efficiency. Overall, 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 on the net.

In the last part of the paper, we 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 the GDP is an added bonus that can lead to further welfare gains through reduced reliance in other sources of taxation that are more distortionary to economic activity. We believe that these results are relevant for industrial policy as well as public finance, and expect future research to delve further into this and other topics relating the interaction between firm intangibles with competition, growth, and social welfare.

References

AGHION, P., A. BERGEAUD, T. BOPPART, P. J. KLENOW, AND H. LI (2019): "A Theory of Falling Growth and Rising Rents," *NBER Working Paper No. 26448*.

AKCIGIT, U., AND S. T. ATES (2019a): "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory," *American Economic Journal: Macroeconomics, forthcoming*.

——— (2019b): "What Happened to US Business Dynamism?," *NBER Working Paper No. 25756*.

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.

BAQAEE, D. R., AND E. FARHI (2020): "Productivity and misallocation in general equilibrium," *The Quarterly Journal of Economics*, 135(1), 105–163.

BARKAI, S. (2020): "Declining Labor and Capital Shares," *The Journal of Finance*, p. jofi.12909.

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*.

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

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, AND X. TIAN (2020): "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, forthcoming*.

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

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.

DeRidder, M. (2020): "Market Power and Innovation in the Intangible Economy," *Working paper*.

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

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

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

Edmond, C., V. Midrigan, and D. Y. Xu (2018): "How Costly Are Markups?," *NBER Working Paper No. 24800*.

Erickson, T., and T. Whited (2002): "Two-step GMM Estimation of the Errors-invariables Model using High-order Moments," *Econometric Theory*, 18, 776–799.

Fitzgerald, D., S. Haller, and Y. Yedid-Levi (2017): "How Exporters Grow," *Working Paper, Federal Reserve Bank of Minneapolis*.

Gourio, F., and L. Rudanko (2014): "Customer Capital," *Review of Economic Studies*, 81(3), 1102 – 1136.

Greenwood, J., Y. Ma, and M. Yorukoglu (2020): "'You Will': A Macroeconomic Analysis of Digital Advertising," *Working paper*.

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*.

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. SENER (2021): "Product innovation, diffusion and endogenous growth," *Working paper*.

LIU, E., A. MIAN, AND A. SUFI (2019): "Low Interest Rates, Market Power, and Productivity Growth," *NBER Working Paper No. 25505*.

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 (2020): "Monetary Policy, Customer Capital, and Market Power," *Working paper*.

PERLA, J. (2019): "Product Awareness, Industry Life Cycles, and Aggregate Profits," *Working Paper, University of British Columbia*.

PETERS, M. (2020): "Heterogeneous Mark-Ups, Growth and Endogenous Misallocation," *Econometrica*, 88(5), 2037–2073.

PETERS, M., AND C. WALSH (2019): "Declining Dynamism, Increasing Markups and Missing Growth: The Role of The Labor Force," *Working paper*.

RACHEL, L. (2021): "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, forthcoming*.

WEISS, J. (2019): "Intangible Investment and Market Concentration," *Working paper.*

Online Appendix:
Style Over Substance?
Advertising, Innovation, and Endogenous Market Structure

A Estimation Details

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 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 simulated moments estimator is defined as the solution to the minimization of weighted average distance between data and model moments. We calculate the optimal weight matrix using influence function method following Erickson and Whited (2002). We use a simulated annealing algorithm to minimize the objective function. This starts with a predefined first and second guess. For the third guess onward, it takes the best prior guess and randomizes from this to generate a new set of parameter guesses. That is, it takes the best-fit parameters and randomly "jumps off" from this point for its next guess. Over time the algorithm "cools", so that the variance of the parameter jumps falls, allowing the estimator to fine-tune its parameter estimates around the global best fit. We restart the program with different initial conditions to ensure the estimator converges to the global minimum. The simulated annealing algorithm is extremely slow, which restricts the size of the parameter space that can be estimated. Nevertheless, we use this because it is robust to the presence of local minima and discontinuities in the objective function across the parameter space.

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.1 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.

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, who report total R&D expenditures divided by GDP.

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 corporate labor share (OECD and UN). For capital share, we rely on the data from Barkai (2020). For both time series, we calculate the averages across all years for our sample. In our 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 (2019b).

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, as in Cavenaile, Celik, and Tian (2020), 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.[17] We normalize both by subtracting their means and dividing by the standard deviation. In the data, we regress average citations on relative sales of the firm in its SIC4 industry and its square. 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.

7. **Average profitability:** In the model, average profitability is calculated as static profit flow minus 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.[18]

10. **Relationship between firm advertising and relative sales:** Replicating the observed inverted-U relationship between competition and advertising also helps use discipline the counterfactual implications of the model regarding 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 (log(1+xad)) on relative sales of the firm in its SIC4 industry and its square. 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

[17]To measure patent citation, we use the patent grant data obtained from NBER Patent Database Project which covers the years 1976-2006. We reply 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).

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

fixed effects. We normalize the advertising expenses in both the data and model by subtracting their means and dividing by the standard deviation, and target the coefficients for the linear and quadratic terms of the regression after normalization.

A.2 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 SMM estimation depends on model identification, which 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 each moment given one percent increase in each parameter. This table gives us some indication on which data moments are most informative in helping us identifying 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 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 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

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 each moment given one percent increase in each parameter.

industry leader and other superstar firms. While 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 inverted-U relationship between innovation and market shares. 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' impacts 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' impacts on linear term of innovation-market share regression are also opposite.

(vii) The exogenous exit rate parameter τ is directly identified by targeting the entry rate of new businesses, since firm entry rate equals 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 growth rate

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

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

where 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 $\Delta \approx 0$, we have:

$$
\begin{aligned}
\ln(Y_{t+\Delta t}) - \ln(Y_t) &= -\ln(w_{t+\Delta t}^{rel}) + \ln(w_t^{rel}) + \ln(1+\lambda) \sum_{\Theta} p_{lit}(\Theta)\Delta t \mu_t(\Theta) \\
&\quad + \sum_{\Theta}\sum_{\Theta'} \left[f_t(\Theta') - f_t(\Theta)\right] p_t(\Theta,\Theta')\mu_t(\Theta)\Delta t + o(\Delta t) \qquad \text{(B.2)}
\end{aligned}
$$

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

$$g_t = -g_{w^{rel},t} + \ln(1+\lambda) \sum_{\Theta} p_{lit}(\Theta)\mu_t(\Theta) + \sum_{\Theta}\sum_{\Theta'} \left[f_t(\Theta') - f_t(\Theta)\right] p_t(\Theta,\Theta')\mu_t(\Theta) \quad \text{(B.3)}$$

B.2 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:

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

B.3 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 initial output, use equation (B.1) to write:

$$Y_0 = \exp\left(\int_0^1 \ln(q_{cj0})dj - \ln(w^{rel}) + \sum_{\Theta} f(\Theta)\mu(\Theta)\right) \quad \text{(B.4)}$$

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

$$C_0 = Y_0 \cdot \frac{C_0}{Y_0} = Y_0 \left(1 + \int_0^1 \sum_{i=1}^{N_{j0}} \chi z_{ij0}^{\phi} \, dj + \int_0^1 \sum_{i=1}^{N_{j0}} \chi_a \omega_{ij0}^{\phi_a} \, dj + \int_0^1 m_0 \nu X_{kj0}^{\epsilon} \, dj + \psi e_0^2 \right) \quad \text{(B.5)}$$

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):

$$W = \int_0^{+\infty} e^{-\rho t} \ln(C_t) \, dt = \frac{1}{\rho} \left(\ln(C_0) + \frac{g}{\rho} \right) \quad \text{(B.6)}$$

Using formulas (B.4) and (B.5), 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:

$$W^B = \frac{1}{\rho} \left(\ln \left(C_0^A (1 + \varsigma) \right) + \frac{g^A}{\rho} \right) \quad \text{(B.7)}$$

Solving for ς, we get:

$$\varsigma = \frac{C_0^B}{C_0^A} \exp \left(\frac{g^B - g^A}{\rho} \right) - 1 \quad \text{(B.8)}$$

C 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.

[19] 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.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:

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

subject to

$$C_t + R_t Y_t + A_t Y_t \leq Y_t \tag{C.2a}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

where equation (C.2a) is the resource constraint; equation (C.2b) is total R&D and business creation investment as a share of GDP; equation (C.2c) is the advertising share of GDP; equations (C.2d) to (C.2j) define production technologies at different levels of aggregation; equation (C.2g) defines the advertising shifter of superstars, equation (C.2k) is the aggregate labor feasibility constraint; equation (C.2l) defines the productivity of the industry leader; equation (C.2m) defines the the productivity of each small firm in the competitive fringe, equation (C.2n) defines the vector of productivity step sizes; equation (C.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 (C.2p) defines the average (log) productivity of leaders across industries; equation (C.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 (C.2r) is the law of motion of the industry distribution; equation (C.2s) states that the mass of industries has to sum to one; and equation (C.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 firms i in industry j at time t (l_{ijt}, ω_{ijt}) and labor to every small firms k in industry j at time t (l_{kcjt}). The social planner also chooses R&D innovation policies for every superstar firms (z_{ijt}) and small firms (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 innovation by small firms 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 (C.2a) and (C.2k)) must bind with equality. Therefore, for a given the 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.

C.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:

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

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

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

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

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

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

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

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

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,

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

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

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

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

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 (C.9), some algebra shows:

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

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

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

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

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

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

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

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

C.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:

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

In equation (C.17), M_t is the aggregate markup defined as a harmonic sales-weighted mean of firm-level markups:[20]

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

Further, one can show that the market shares (defined in equation (20)) can be written in terms of markups as follows:

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

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

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

[20] It is easy to show that the aggregate markup is the inverse of the aggregate labor share, which equals the relative wage $w_t^{rel}\equiv\frac{w_t}{Y_t}$ because the labor force is normalized to one ($L=1$).

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).

D Additional Figures

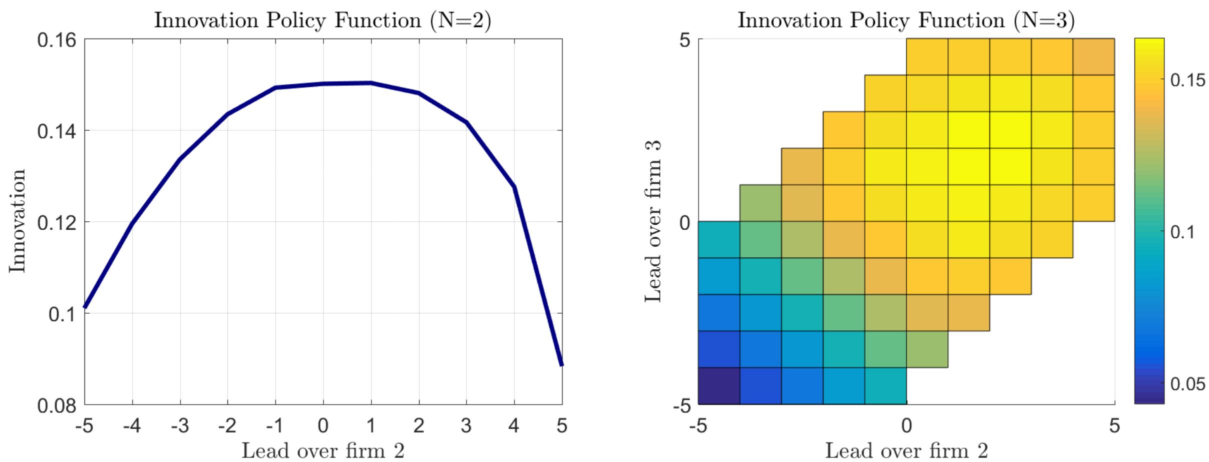

FIGURE D.1: INNOVATION POLICY FUNCTION

Notes: This figure displays the optimal innovation policy functions followed by the firms in the estimated equilibrium. The left panel depicts the innovation policy of a firm in an industry with two superstar firms. The relevant state variable is how many steps the current firm is ahead of its competitor, where negative numbers indicate that the current firm is lagging behind the competing firm. The right panel does the same for a firm in an industry with three superstar firms.