SIAM REVIEW
Vol. 64, No. 3, pp. 625–639

A Mathematical Model for the Origin of Name Brands and Generics*

Joseph D. Johnson†
Adam M. Redlich†
Daniel M. Abrams‡

Abstract. Firms in the U.S. spend over $200 billion each year advertising their products to consumers, around one percent of the country's gross domestic product. It is of great interest to understand how that aggregate expenditure affects prices, market efficiency, and overall welfare. Here, we present a mathematical model for the dynamics of competition through advertising and find a surprising prediction: when advertising is relatively cheap compared to the maximum benefit advertising offers, rational firms split into two groups, one with significantly less advertising (a "generic" group) and one with significantly more advertising (a "name-brand" group). Our model predicts that this segmentation will also be reflected in price distributions; we use large consumer data sets to test this prediction and find good qualitative agreement.

Key words. dynamical systems, nonlinear dynamics, differential games, advertising, economic dynamics, consumer behavior

AMS subject classifications. 35Q91, 37N40, 91-10, 91A16, 91A23, 91B15, 91B42, 91B55

DOI. 10.1137/20M1360888

1. Introduction and Background.

Advertising is an important component of a free market system; it has been estimated that advertising expenditures in the United States exceeded $200 billion in 2018 alone [32]. Although the monetary investment is large, it remains unclear exactly how advertising affects demand and what the implications are for market competition. Perhaps advertising leads to increased market efficiency, greater aggregate profit for sellers, or better outcomes for buyers. The opposite could also be argued.

There are three prevailing theories as to how advertising influences the consumer [30]. Advertising can be viewed as *persuasive*, whereby it changes the tastes of consumers and increases demand (and price) [5, 11, 12]; *informative*, whereby it increases competition and decreases price [36, 26, 34]; or *complementary*, whereby it appeals to consumers with specific preferences that complement the consumption of the advertised products [4, 27, 35]. These views have drastically different implications.

*Received by the editors August 19, 2020; accepted for publication (in revised form) September 20, 2021; published electronically August 4, 2022.
https://doi.org/10.1137/20M1360888
Funding: This work was supported by the National Science Foundation through Research Training grant 1547394.
†Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, IL 60208 USA (josephjohnson2020@u.northwestern.edu, adamredlich2020@u.northwestern.edu).
‡Department of Engineering Sciences and Applied Mathematics, Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL 60208 USA (dmabrams@northwestern.edu).

In this paper we focus on *persuasive* advertising and, as in [6], assume that it increases demand. We look to work by Abernethy and Butler [1] to justify this assumption, where they report that an average TV ad contains just one piece of descriptive information about the displayed product (e.g., price, quality, performance, etc.), and that 37.5 percent contain no descriptive information at all. We take this to mean that a significant portion of TV ads are not informative, implying that they are persuasive or complementary. Additionally, we make the simplifying assumption that persuasive advertising is always complementary, as Lindstädt and Budzinski argue that the viewer relates to the images and messages for both complementary and persuasive advertising [21].

A large amount of research has been devoted to using game theory to choose the optimal advertising expenditure to maximize profit [10, 22, 33, 31, 16, 13, 15]. Often this work focuses on settings where there is a monopoly (only one supplier of a good or service) or an oligopoly (only a small number of suppliers of a good or service) [14, 13, 31, 16, 15, 22, 17].

Less research has focused on *monopolistic competition*, where there are many suppliers of a product or service, but the products or services are differentiated only by brand and/or quality. In this paper we develop a model for this setting, looking at the expected advertising expenditure distribution for an arbitrary number of firms competing in a single commodity-product sector. Our goal is to develop a qualitative understanding of the expected shape of the advertising distribution in a monopolistic competitive setting.

1.1. Synopsis of Modeling Approach. In developing our model, we make the following simplifying assumptions:

1. Companies[1] sell an indistinguishable product (except for brand label).
2. There is a linear relationship between the amount of a company's product demanded by the public and the price of the product.
3. Demand for a company's product increases when its advertising is above the mean advertising level and decreases when its advertising is below the mean.
4. Each company sets the price at a level that maximizes its profit.
5. Companies continuously adjust their advertising so as to maximize profit.

These assumptions lead to a system of ordinary differential equations describing the dynamics of advertising investments for N firms.

These equations imply that, when advertising is relatively cheap compared to the benefit of advertising, two groups arise: a "generic-brand" group that advertises a minimal amount, and a "name-brand" group that advertises at a significantly higher level.[2] We find that this segmentation is stable and only ceases to exist when the marginal cost of advertising becomes too high relative to the marginal benefit of advertising. Although our model is intended chiefly to provide a conceptual "toy" description, fits to real-world price data[3] show good qualitative agreement (see Figure 1).

We caution the reader that, though we use the terms "generic" and "name brand"

[1] We use the terms "companies" and "firms" interchangeably.

[2] Note that the minimal advertising level may not be zero: there may be some fixed advertising costs, e.g., associated with product packaging or distribution. We will treat the minimal level as zero (representing zero "excess" advertising) for simplicity in presenting our model, but including an additive constant does not change our predictions.

[3] Data in Figure 1 have been treated to compensate for psychological pricing, where prices tend to end in certain digits such as "0", "5", and "9". See the supplementary material for details.

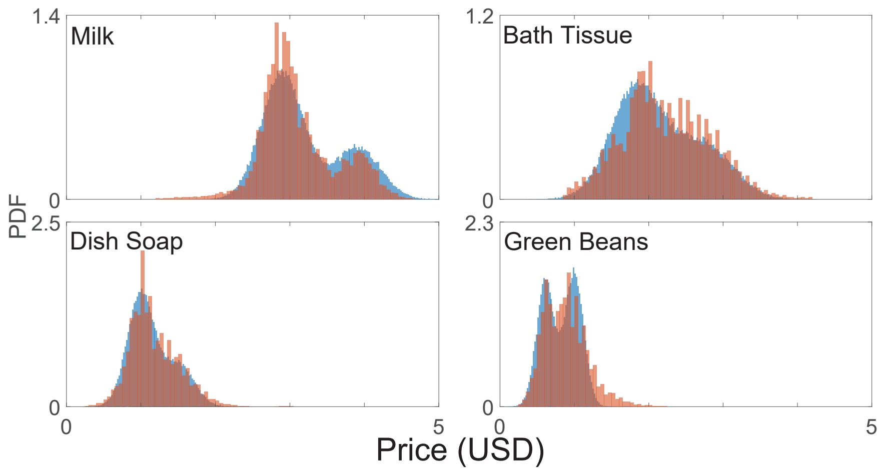

Fig. 1 Price distributions. *Red histograms show distribution of prices paid for four common household products; blue histograms show best-fit model predictions. Data are from a Nielsen database* [7] *of over* 64 *million transactions (purchasing history for* 60,000 *households).*

to refer to low and high advertising investment states, this is an oversimplification. In the real world, some brands considered generic may in fact spend significantly on advertising, and some well-known name brands may invest very little in it. However, for simplicity of exposition, we will employ these terms throughout the paper.

1.2. Outline. In section 2, we present our mathematical model for how firms work to differentiate themselves through advertising. We also present our results relating to existence and stability of equilibria. In section 3, we report briefly on the results of numerical experiments to verify consistency with model predictions. In section 4, we present real-world data on price distributions for a variety of products and evaluate model fits to assess the consistency of our predictions. Finally, in section 5 we discuss other possible applications of our model and limitations of our results.

2. Model and Analysis.

2.1. Model Derivation. Consider N companies (or firms) in a market, all selling the same indistinguishable[4] product. The ith firm purchases a quantity of advertising a_i.[5] For simplicity we assume that the firms have linear demand curves of the form

$$Q_i = Q_{\text{free}}(a_i|\vec{a}) - k_{\text{P}} P_i \ , \qquad i = 1, 2, \ldots, N, \tag{2.1}$$

where Q_i is the quantity demanded of firm i's product, P_i is the unit price for firm i's product, $Q_{\text{free}}(a_i|\vec{a})$ is the quantity demanded when the unit price is zero, which may depend on the full distribution of advertising in the market $\vec{a} = (a_1, a_2, \ldots, a_N)$, and k_{P} is a constant that sets the market's sensitivity to price.

One measure of a firm's health is the profit generated, with profit defined here as revenue minus production and advertising costs. We take revenue R_i for the ith firm

[4]By "indistinguishable" we mean that the product without branding is indistinguishable, but the brand label is always known to the consumer.

[5]This could be quantified, e.g., by clicks on a website ad banner, inserts in a newspaper, views of an ad on TV, or supermarket placement costs.

to be solely due to sales of this single product at market price:

$$R_i = Q_i P_i = Q_{\text{free}}(a_i|\vec{a})P_i - k_{\text{P}} P_i^2, \quad i = 1, 2, \ldots, N. \tag{2.2}$$

In this model we only consider two types of operating costs, the cost of production $C_{\text{Q}}(Q_i)$ and the cost of advertising $C_{\text{a}}(a_i)$, and we assume an additive relationship

$$C_i(Q_i, a_i) = C_{\text{Q}}(Q_i) + C_{\text{a}}(a_i), \tag{2.3}$$

where $C_i(Q_i, a_i)$ is the net operating cost for the ith firm. We assume that both C_{Q} and C_{a} are increasing functions of their arguments, and for simplicity[6] assume a power law form for each:

$$C_{\text{Q}}(Q_i) = k_{\text{Q}} Q_i^{\mu}, \tag{2.4a}$$

$$C_{\text{a}}(a_i) = k_{\text{a}} a_i^{\nu}, \tag{2.4b}$$

where $\mu, \nu > 0$ and k_{Q}, k_{a} are scale factors and can be interpreted as the marginal costs of production and advertising, respectively, when $\mu = \nu = 1$. Thus, the profit function for the ith firm is

$$\pi_i = R_i - C_i = Q_{\text{free}}(a_i|\vec{a})P_i - k_{\text{P}} P_i^2 - k_{\text{Q}} Q_i^{\mu} - k_{\text{a}} a_i^{\nu}. \tag{2.5}$$

Critically, we tie a firm's level of advertising, a_i, to its ability to capture market power. We do this by assuming $Q_{\text{free}}(a_i|\vec{a})$ to be a nondecreasing function of a_i referenced to the mean advertising level $\overline{a} = N^{-1}\Sigma_{i=1}^{N} a_i$, i.e., a nondecreasing function of $a_i - \overline{a}$ (in the most general case, however, it might be an arbitrary function of the full advertising distribution $\vec{a} = (a_1, \ldots, a_N)$). We assume firms that advertise more than the average firm have their demand curves shift out (i.e., quantity demand increases by a constant amount for all P_i) and firms that advertise less than average have their demand curves shift in (i.e., quantity demand decreases by a constant amount for all P_i)—see Figure 2.

We also assume there is a saturation to the amount advertising can influence a firm's ability to capture market share. A plausible smooth, nondecreasing function that saturates is the sigmoid. We present results for that case in the supplementary material (SM). For greater algebraic simplicity, we define $Q_{\text{free}}(a_i|\vec{a})$ here as the following saturating piecewise linear function:

$$Q_{\text{free}}(a_i|\vec{a}) = \begin{cases} Q_{\min}, & a_i - \overline{a} \leq -\lambda\,, \\ Q_{\min} + \dfrac{\Delta Q_{\text{ad}}}{2\lambda}(a_i - \overline{a}) + \dfrac{\Delta Q_{\text{ad}}}{2}, & -\lambda < a_i - \overline{a} \leq \lambda\,, \\ Q_{\min} + \Delta Q_{\text{ad}}, & a_i - \overline{a} > \lambda\,, \end{cases} \tag{2.6}$$

where ΔQ_{ad} is the maximum demand increase due to advertising, $Q_{\min}$ is the zero-advertising (minimum) quantity demanded at zero price, which we deem "intrinsic demand," and λ is the width of $Q_{\text{free}}(a_i|\vec{a})$ (roughly the amount of excess advertising—above or below the mean—needed for benefits to saturate). See Figure 2 for an illustration. Note, however, that for the purpose of comparison with data, we use the more plausible sigmoidal form

$$Q_{\text{free}}(a_i|\vec{a}) = \frac{\Delta Q_{\text{ad}}}{2}\left\{\tanh\left[\frac{a_i - \overline{a}}{\lambda}\right] + 1\right\} + Q_{\min}. \tag{2.7}$$

[6]Power laws are common in both natural and engineered systems [23, 8], and there is evidence that production costs can indeed be approximated by power law scaling [38].

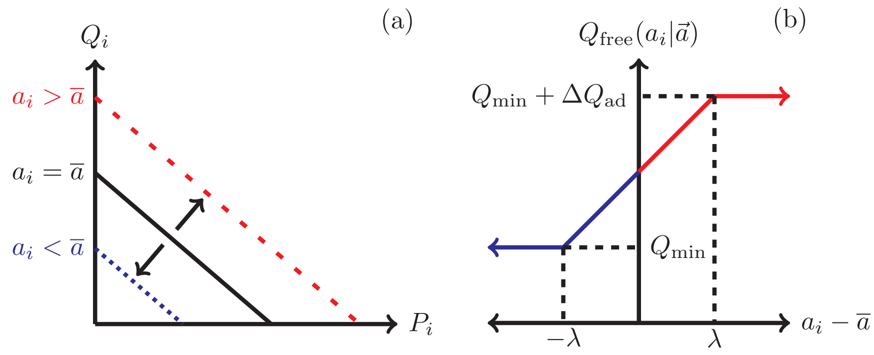

Fig. 2 Effect of advertising on a firm's demand curve. (a) *Demand shifts due to advertising above (red dashed) or below (blue dotted) the mean level (black solid). Vertical-axis intercepts are* $Q_{\mathrm{free}}(a_i|\vec{a})$. (b) *A simple piecewise linear form for* $Q_{\mathrm{free}}(a_i|\vec{a})$, *the quantity demanded at zero price, which we take to be a nondecreasing function of* $a_i - \overline{a}$ *that saturates at both left and right limits. Here the minimum demand (with advertising far below the mean) is* $Q_{\min}$, *the maximum demand increase due to advertising is* ΔQ_{ad}, *and the advertising needed beyond the mean for saturation is* λ.

Table 1 Parameter definitions. *Table of parameters used in the model with descriptions.*

Parameter	Description
N	Number of companies
$Q_{\min}$	The quantity demanded with minimal advertising at zero price
ΔQ_{ad}	The maximum demand increase due to advertising
k_{P}	Decrease in quantity demanded per dollar in unit price increase
k_{Q}	Scale factor for production cost (cost of producing an additional unit when costs are linear)
k_{a}	Scale factor for advertising cost (cost of producing an additional advertisement when costs are linear)
λ	Amount of excess advertising above/below the mean to achieve maximum/minimum advertising benefits
μ	Scaling exponent in the production cost function
ν	Scaling exponent in the advertising cost function

We assume that each firm always chooses the price P_i^* that maximizes its profit, with corresponding quantity demanded Q_i^*. We introduce dynamics to the model by assuming that firms change their advertising levels at a rate proportional to the amount of profit to be gained, i.e.,

(2.8)
$$\tau \frac{da_i}{dt} = \frac{\partial \pi_i}{\partial a_i} = \frac{\partial}{\partial a_i} \left\{ Q_{\mathrm{free}}(a_i|\vec{a}) P_i^*(a_i|\vec{a}) - k_{\mathrm{P}} [P_i^*(a_i|\vec{a})]^2 - k_{\mathrm{Q}} Q_i^*(a_i|\vec{a})^{\mu} - k_{\mathrm{a}} a_i^{\nu} \right\},$$

where the constant τ sets the time scale for equilibration; we will henceforth take $\tau = 1$ (equivalent to rescaling the time axis) without loss of generality. The list of model parameters with definitions is given in Table 1.

2.2. A Concrete Example. As an analytically tractable example, we first consider the case where production and advertising costs grow at a linear rate, i.e., $\mu = \nu = 1$. Substituting (2.1) into (2.5), setting $[\partial \pi_i / \partial P_i]_{P_i = P_i^*} = 0$, and solving for

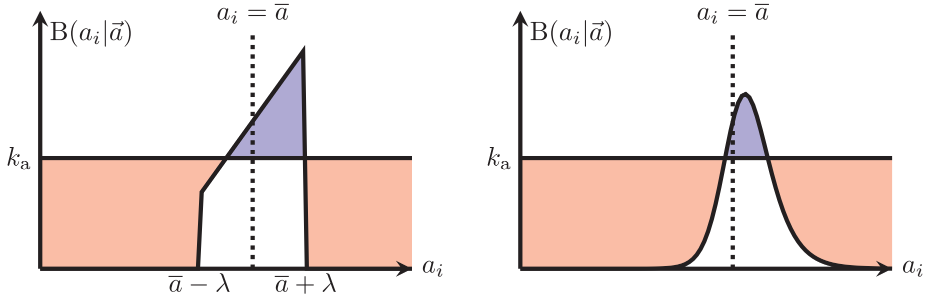

Fig. 3 Advertising dynamics given $Q_{\mathrm{free}}(a_i|\vec{a})$. *We plot the the horizontal line* $y = k_{\mathrm{a}}$ *together with* $y = \mathrm{B}(a_i|\vec{a})$ *(see* (2.12)*) and add color to indicate the direction of change in advertising according to* (2.8). *Blue regions show where a firm advertising an amount* a_i *would choose to increase its advertising, and red regions show where a firm advertising an amount* a_i *would choose to decrease its advertising. Panel* (a) *shows the case where* $Q_{\mathrm{free}}(a_i|\vec{a})$ *is a piecewise linear function that levels off;* (b) *shows the case where* $Q_{\mathrm{free}}(a_i|\vec{a})$ *is sigmoidal.*

the profit-maximizing price P_i^* gives

$$P_i^*(a_i|\vec{a}) = \frac{1}{2}\left[Q_{\mathrm{free}}(a_i|\vec{a})/k_{\mathrm{P}} + k_{\mathrm{Q}}\right]. \tag{2.9}$$

The corresponding profit-maximizing quantity is

$$Q_i^*(a_i|\vec{a}) = \frac{1}{2}\left[Q_{\mathrm{free}}(a_i|\vec{a}) - k_{\mathrm{Q}}k_{\mathrm{P}}\right]. \tag{2.10}$$

Substituting this into (2.8) yields the dynamical system

$$\frac{da_i}{dt} = \mathrm{B}(a_i|\vec{a}) - k_{\mathrm{a}}, \tag{2.11}$$

where $\mathrm{B}(a_i|\vec{a})$ is defined as

$$\mathrm{B}(a_i|\vec{a}) = \begin{cases} \dfrac{N-1}{N}\dfrac{\Delta Q_{\mathrm{ad}}}{4\lambda k_{\mathrm{P}}}\left[\dfrac{\Delta Q_{\mathrm{ad}}}{2\lambda}(a_i - \overline{a}) + \dfrac{\Delta Q_{\mathrm{ad}}}{2} + Q_{\mathrm{min}} - k_{\mathrm{Q}}k_{\mathrm{P}}\right], & |a_i - \overline{a}| < \lambda, \\ 0, & |a_i - \overline{a}| > \lambda. \end{cases} \tag{2.12}$$

$\mathrm{B}(a_i|\vec{a})$ represents the marginal benefit of advertising and k_{a} the marginal cost of advertising. For any firm with advertising close enough to the mean ($|a_i - \overline{a}| < \lambda$), the function B is simply a line of positive slope $(N-1)\Delta Q_{\mathrm{ad}}^2/(8N\lambda^2 k_{\mathrm{P}}) \xrightarrow[N\to\infty]{} \Delta Q_{\mathrm{ad}}^2/8\lambda^2 k_{\mathrm{P}}$. Firms with $\mathrm{B} > k_{\mathrm{a}}$ have $da_i/dt > 0$ and increase their advertising budgets, while firms with $\mathrm{B} < k_{\mathrm{a}}$ decrease their advertising budgets. For all firms far from the mean ($|a_i - \overline{a}| > \lambda$), $\mathrm{B} = 0$ and thus $da_i/dt = -k_{\mathrm{a}} < 0$. This flow is illustrated in the left panel of Figure 3. The corresponding flow in the case of a smooth sigmoidal $Q_{\mathrm{free}}(a_i|\vec{a})$ is shown in the right panel of the same figure. The intuition drawn from the piecewise case outlined here applies similarly to the sigmoid.

2.2.1. Existence of Equilibria. For a given $\overline{a}$, there can be at most three fixed points. In Figure 3, three fixed points are located at $a_i^* = \overline{a} + \lambda$, the intersection where $\mathrm{B}(a_i^*|\overline{a}) = k_{\mathrm{a}}$ for $a_i^* < \overline{a}$, and at $a_i^* = 0$ (since advertising cannot be negative).

Because stability must alternate for one-dimensional flows, $a_i^* = 0$ and $a_i^* = \overline{a} + \lambda$ are the only stable fixed points. Thus, any stable equilibrium distribution $\vec{a}$ with mean $\overline{a}$ must have $a_i = 0$ or $a_i = \overline{a} + \lambda$ for all i.

We refer to the case when advertising is bimodal as the *differentiated* state. We note that such a state may only exist when two stable fixed points exist, which requires $\max_a \mathrm{B}(a|\vec{a}) > k_\mathrm{a}$. Letting $N \to \infty$, one can write this condition explicitly as

$$\max_{a_i} \frac{\partial \pi_i}{\partial a_i} = \frac{\Delta Q_\mathrm{ad}\left(\Delta Q_\mathrm{ad} + Q_\mathrm{min} - k_\mathrm{Q} k_\mathrm{P}\right)}{4\lambda k_\mathrm{P}} - k_\mathrm{a} > 0. \tag{2.13}$$

Put simply, if advertising does not increase profit anywhere, bimodality cannot arise.

The differentiated state must be self-consistent—that is, the two stable fixed points (at 0 and $\overline{a}+\lambda$) when averaged (including weights based on corresponding firm fractions) must yield the appropriate mean advertising $\overline{a}$. We refer to the fraction of firms that choose to set their advertising to zero ("generics") as x, and thus the fraction of firms that set their advertising to $a_i^* = \overline{a} + \lambda$ ("name brands") is $1 - x$. It follows then that $\overline{a} = 0x + a_i^*(1-x) = (\overline{a} + \lambda)(1-x)$ must hold for self-consistency. Solving for $\overline{a}$ gives

$$\overline{a} = \frac{\lambda(1-x)}{x} \ , \tag{2.14}$$

and the name-brand advertising level at equilibrium is then $a_i^* = \overline{a} + \lambda = \lambda/x$.

We also require that the unstable fixed point at $\mathrm{B}(a_i^*|\overline{a}) = k_\mathrm{a}$ be nonnegative, assuming the system has reached self-consistent equilibrium (i.e., $\overline{a} = \lambda(1-x)/x$ as in (2.14)). If the unstable fixed point were negative, that would imply $da_i/dt > 0$ over the entire domain $0 \le a_i < \lambda/x$, which would contradict the assumption of a stable fixed point at $a_i = 0$ that went into the self-consistency argument above.

Imposing this constraint on the unstable fixed point defined by $\mathrm{B}(a_i^*|\overline{a}) = k_\mathrm{a}$, we find

$$a_i^{*(\mathrm{unstable})} = \frac{\lambda}{x\Delta Q_\mathrm{ad}^2}\left[\Delta Q_\mathrm{ad}^2(1-2x) - 2x\Delta Q_\mathrm{ad}(Q_\mathrm{min} - k_\mathrm{P} k_\mathrm{Q}) + 8x\lambda\left(\frac{N}{N-1}\right) k_\mathrm{a} k_\mathrm{P}\right] \tag{2.15}$$

and

$$x_\mathrm{crit} = \frac{\Delta Q_\mathrm{ad}^2/2}{\Delta Q_\mathrm{ad}^2 + \Delta Q_\mathrm{ad}(Q_\mathrm{min} - k_\mathrm{P} k_\mathrm{Q}) - 4\lambda\left(\frac{N}{N-1}\right) k_\mathrm{P} k_\mathrm{a}}. \tag{2.16}$$

Here, x_crit bounds the feasible proportion of generic firms from above. Equations (2.13) and (2.16) (with $x < x_\mathrm{crit}$) establish necessary conditions for existence of the differentiated state. It is not feasible to derive an equivalent analytical expression for x_crit in the case of sigmoidal $Q_\mathrm{free}(a_i|\vec{a})$, but it is straightforward to compute x_crit numerically for a given set of parameters, as we do in the SM.

Another state is possible where all firms set their advertising to zero—we refer to this as the *undifferentiated* state. Clearly, from (2.11), $\max_a \mathrm{B}(a|\vec{a}) < k_\mathrm{a}$ implies that $da_i/dt < 0$ for all a_i. In this case, $a_i^* = 0$ for all i is the only equilibrium.

2.2.2. Stability of Equilibria. We now consider the stability of the differentiated and undifferentiated states. First, we focus on the stability of the differentiated state. We assume there exists an equilibrium with Nx "generic" firms choosing to invest nothing in advertising, and $N(1-x)$ "name-brand" firms choosing to advertise at level

$\overline{a}+\lambda$, with $0<x<1$ representing the proportion of "generic" firms. Assuming that $N \gg 1$[7] and hence that a small perturbation of a single firm has a negligible impact on the mean $\overline{a}$, we consider perturbation of the ith "name-brand" firm's advertising by an amount δ and track how $\delta(t)$ changes in time. That is, we set $a_i = \overline{a}+\lambda+\delta(t)$, which yields the system

$$\frac{d\delta}{dt} = \begin{cases} \frac{N-1}{N}\frac{\Delta Q_{\mathrm{ad}}}{4\lambda k_{\mathrm{P}}}\left[\frac{\Delta Q_{\mathrm{ad}}}{2\lambda}\delta + \Delta Q_{\mathrm{ad}} + Q_{\mathrm{min}} - k_{\mathrm{Q}}k_{\mathrm{P}}\right] - k_{\mathrm{a}}, & |\delta+\lambda|<\lambda, \\ -k_{\mathrm{a}}, & |\delta+\lambda|>\lambda. \end{cases} \tag{2.17}$$

If the condition for existence of the differentiated state given in (2.13) holds, sufficiently small $|\delta|$ implies that $d\delta/dt > 0$ when $\delta < 0$. Additionally, it is clear that $d\delta/dt < 0$ when $\delta > 0$. Thus, under this type of perturbation the differentiated state is stable. If we similarly perturb one firm from the generic group, i.e., setting $a_i = \delta > 0$, we find

$$\frac{d\delta}{dt} = \begin{cases} \frac{N-1}{N}\frac{\Delta Q_{\mathrm{ad}}}{4\lambda k_{\mathrm{P}}}\left[\frac{\Delta Q_{\mathrm{ad}}}{2\lambda}(\delta-\overline{a}) + \frac{\Delta Q_{\mathrm{ad}}}{2} + Q_{\mathrm{min}} - k_{\mathrm{Q}}k_{\mathrm{P}}\right] - k_{\mathrm{a}}, & |\delta-\overline{a}|<\lambda, \\ -k_{\mathrm{a}}, & |\delta-\overline{a}|>\lambda. \end{cases} \tag{2.18}$$

If $\overline{a} > \lambda$, then there exists $\delta > 0$ small enough such that $d\delta/dt < 0$ since $\delta < \overline{a}+\lambda$ implies that $d\delta/dt = -k_{\mathrm{a}} < 0$. If $\overline{a} < \lambda$, then $d\delta/dt$ is given by the linear equation in (2.18) for small δ. Thus, the differentiated state is stable under such a perturbation when

$$\left.\frac{d\delta}{dt}\right|_{\delta\to 0^+} = \frac{N-1}{N}\frac{\Delta Q_{\mathrm{ad}}}{4\lambda k_{\mathrm{P}}}\left[-\frac{\Delta Q_{\mathrm{ad}}}{2\lambda}\overline{a} + \frac{\Delta Q_{\mathrm{ad}}}{2} + Q_{\mathrm{min}} - k_{\mathrm{Q}}k_{\mathrm{P}}\right] - k_{\mathrm{a}} = \left.\frac{\partial \pi_i}{\partial a_i}\right|_{a_i=0} < 0. \tag{2.19}$$

This means it must be unprofitable for companies with no advertising to increase their advertising for the differentiated state to be stable.

Now we consider the stability of the undifferentiated state, $a_i = 0$ for all i. As stated in section 2.2.1, if $\max_a \mathrm{B}(a|\vec{a}) < k_{\mathrm{a}}$, then $da_i/dt < 0$ for all values of a_i. Thus, it is clear that the undifferentiated state exists and is stable in that case. We now focus on the case where $\max_a \mathrm{B}(a|\vec{a}) > k_{\mathrm{a}}$. If $a_i = 0$ for all i, then $\overline{a} = 0$. We consider a perturbation of one firm from this state. Letting $a_i = \delta$, again we find

$$\frac{d\delta}{dt} = \begin{cases} \frac{N-1}{N}\frac{\Delta Q_{\mathrm{ad}}}{4\lambda k_{\mathrm{P}}}\left[\frac{\Delta Q_{\mathrm{ad}}}{2\lambda}(\delta) + \frac{\Delta Q_{\mathrm{ad}}}{2} + Q_{\mathrm{min}} - k_{\mathrm{Q}}k_{\mathrm{P}}\right] - k_{\mathrm{a}}, & 0 \le \delta < \lambda, \\ -k_{\mathrm{a}}, & \delta > \lambda. \end{cases} \tag{2.20}$$

If $d\delta/dt < 0$ when $\delta = 0$, by continuity of the $d\delta/dt$ in the range of $0 < \delta < \lambda$ there must exist some $\delta > 0$ sufficiently small such that $d\delta/dt < 0$. Therefore, the system is stable under this kind of perturbation when

$$\left.\frac{d\delta}{dt}\right|_{\delta\to 0} = \frac{N-1}{N}\frac{\Delta Q_{\mathrm{ad}}[Q_{\mathrm{min}} + \Delta Q_{\mathrm{ad}}/2 - k_{\mathrm{Q}}k_{\mathrm{P}}]}{4\lambda k_{\mathrm{P}}} - k_{\mathrm{a}} = \left.\frac{\partial \pi_i}{\partial a_i}\right|_{a_i=\overline{a}} < 0. \tag{2.21}$$

We surmise from this that the undifferentiated state is stable only if increasing advertising is not profitable for the average firm.

[7]Numerical experiments suggest that stability conditions derived in this section also hold for small N.

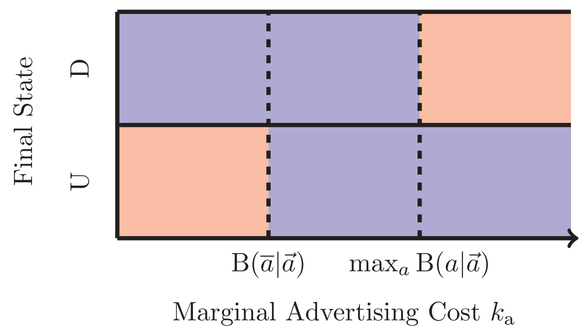

Fig. 4 Regions of stability. *We illustrate the regions of stability for the differentiated and undifferentiated states, indicated in the figure by D and U, respectively. These are given by* (2.13), (2.21), *and* (2.22). *Here blue indicates that a state is stable and red indicates that a state is unstable. The middle column, where* $\mathrm{B}(\overline{a}|\vec{a}) < k_\mathrm{a} < \max_a \mathrm{B}(a|\vec{a})$*, is where both states are stable.*

The above stability arguments can be generalized to arbitrary infinitesimal perturbations of the advertising distribution in the limit $N \to \infty$. See work by Clifton, Braun, and Abrams for a description of such an approach in a different context [9].

Figure 4 maps the regions of stability for the differentiated and undifferentiated states given by (2.13), (2.21), and (2.22). Both the differentiated and undifferentiated states can be simultaneously stable. If (2.13) and (2.21) both hold, then both states are stable. Thus, we write the condition for bistability as

$$\mathrm{B}(\overline{a}|\vec{a}) = \frac{N-1}{N}\frac{\Delta Q_\mathrm{ad}[Q_\mathrm{min} + \Delta Q_\mathrm{ad}/2 - k_\mathrm{Q}k_\mathrm{P}]}{4\lambda k_\mathrm{P}} < k_\mathrm{a} < \max_a \mathrm{B}(a|\vec{a}). \tag{2.22}$$

In more intuitive terms,

$$\left.\frac{\partial \pi_i}{\partial a_i}\right|_{a_i=\overline{a}} < 0 < \max_a \frac{\partial \pi_i}{\partial a_i}. \tag{2.23}$$

Thus, bistability of the differentiated and the undifferentiated states occurs when the maximum marginal profit is positive, but it is profitable for the average firm to decrease its advertising. The regions of stability of the undifferentiated and differentiated states are defined similarly when $Q_\mathrm{free}(a_i|\vec{a})$ is sigmoidal (and, hence, $\mathrm{B}(a_i|\vec{a})$ altered appropriately—see the SM).

3. Numerical Experiments.

In order to test model predictions, we perform simple numerical experiments; all results appear to be consistent with theory. Figure 5 shows an example of a simulation where the benefit of advertising saturates (we assume a sigmoidal functional form) when $k_\mathrm{a} < \max_a \mathrm{B}(a|\vec{a})$. Starting from a uniformly distributed initial condition, the firms arrange themselves so that there is a "generic" group at advertising level $a = 0$ and a "name-brand" group at $a = a_\mathrm{name} > 0$. Colors have been added to indicate ranges where firms decrease (red) or increase (yellow) their advertising (see figure caption for details).

Figure 6 demonstrates some of the existence and stability boundaries outlined in section 2. Panels (a) and (b) start with an initial condition that is sampled from the uniform random distribution $\mathcal{U}(7.5, 12.5)$. When $k_\mathrm{a} < \max_a \mathrm{B}(a|\vec{a})$ (panel (a)),

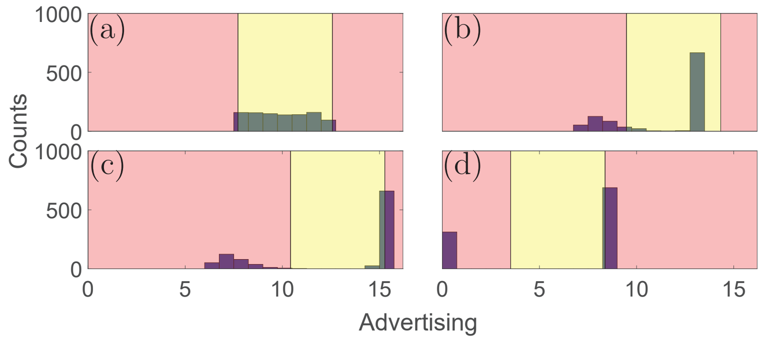

Fig. 5 Simulation of the system. *In this figure we give snapshots of the numerical integration of the system from initial condition to equilibrium. In panel* (a)*, the system starts from the uniform randomly distributed state with the advertising initial condition set as* $\mathcal{U}(7.5, 12.5)$*. In panel* (b)*, the separation into two groups has begun. Companies change their spending until the lower group is far away from the mean, as seen in panel* (c)*. Finally, in panel* (d)*, a bimodal equilibrium has been reached, with one group representing generic brand companies* ($a \approx 0$) *and the other representing name-brand companies (advertising at a nonzero value at* $a = a_{name}$*). The green areas indicate where companies will increase their advertising and the red areas indicate areas where companies will decrease their advertising. In this simulation we set the number of companies to* $N = 1000$, $k_{\mathrm{a}} = k_{\mathrm{P}} = k_{\mathrm{Q}} = \nu = \mu = \lambda = 1$, *and* $Q_{\min} = \Delta Q_{ad} = 10$ *(see Table* 1 *for parameter definitions).*

firms separate into two groups and move toward the differentiated state. When $k_{\mathrm{a}} > \max_a \mathrm{B}(a|\vec{a})$ (panel (b)), all firms tend to zero advertising.

Panels (c) and (d) each begin with bimodal advertising distributions, differing only in the initial fractions generic x. In panel (c), initially $x < x_{\mathrm{crit}}$, and the system relaxes to a stable differentiated state. In panel (d), initially $x > x_{\mathrm{crit}}$, and, since no nearby differentiated state exists for that fractionation, the system ends up at a different differentiated state (where $x < x_{\mathrm{crit}}$) due to some firms transitioning from the generic group to the name-brand group. See figure caption for more detail.

In the SM, we discuss simulation of other variants of our model including nonlinear production cost curves ($\mu \neq 1$) and advertising cost curves ($\nu \neq 1$) and nonidentical firms (e.g., nonuniform $Q_{\min}$ and/or ΔQ_{ad}). We encountered no qualitative difference in the results for those cases.

4. Data. We use price data from the Nielsen Corporation. Nielsen's consumer panel data contains annual shopping information from thousands of American households, starting from 2004 with yearly updates. Individuals involved in the study used in-home scanners to record all of their purchases that were designated for personal use. Scanners recorded each product's Universal Product Code (a string of digits that uniquely identify the product) and the product's price. We analyze data from 2014 containing over 64 million transactions from 60,000 households [7].

Our model's primary prediction is the distribution of advertising investments across firms. While we would have preferred to employ data that directly reflects such advertising budgets, we were not able to find any source comparable in quality to the Nielsen price data set. Nevertheless, our model also carries with it predictions for prices, though we must accept that real-world prices may be additionally influenced

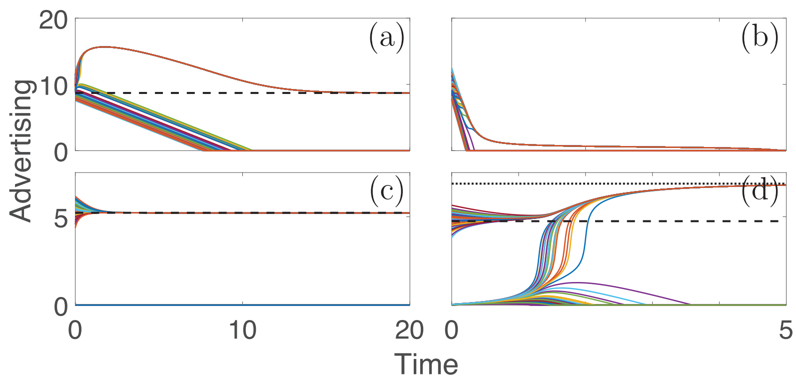

Fig. 6 Numerical exploration. *We illustrate the qualitative states described in section* 2. *In panels* (a) *and* (b)*, we set the initial condition by sampling from the uniform random distribution* $\mathcal{U}(7.5, 12.5)$. *In panel* (a), $k_\mathrm{a} < \max_a \mathrm{B}(a|\vec{a})$, *and the firms approach a bimodal distribution (differentiated state). In panel* (b), $k_\mathrm{a} > \max_a \mathrm{B}(a|\vec{a})$, *and the firms approach the zero advertising (undifferentiated) state. In panels* (c) *and* (d)*, we set the initial condition by perturbing off a theoretical differentiated state with* $x_{crit} = 0.53$. *In panel* (c), *we set* $x = 0.5 < x_{crit}$*; after perturbation the system returns to the differentiated state. In panel* (d), *we set* $x = 0.55 > x_{crit}$*; after perturbation some firms move from the generic group to the name-brand group so that the final generic fraction* $x = 0.38$ *is less than* $x_{crit} = 0.53$. *Dashed lines show the theoretical advertising level for the name-brand group before initial perturbation. The dotted line in panel* (d) *shows the theoretical advertising level for the name-brand group after the generic fraction has changed to its final value. In all simulations we set* $N = 100$, $k_\mathrm{P} = k_\mathrm{Q} = \nu = \mu = \lambda = 1$, *and* $Q_\mathrm{min} = \Delta Q_{ad} = 10$.

by other unmodeled factors. In working with price data, we make the assumption that these other unmodeled factors either have negligible impact or do not change the unimodal/multimodal nature of the distribution.

4.1. Fitting Procedure. To fit our model predictions to data, we first define an objective function $H[f(p), g(p)]$ to quantify the difference between distributions predicted by the model ($f(p)$) and inferred from the data ($g(p)$). Specifically, we set our objective function $H[f(p), g(p)]$ to be the square integrated difference between the distributions

$$H(f, g) = \int_{-\infty}^{\infty} [f(p) - g(p)]^2 dp. \tag{4.1}$$

We use the Nelder–Mead algorithm [25] to minimize this objective function over a subset of parameters that most directly affect the demand curve given in (2.1): the maximum benefit from advertising ΔQ_ad, the minimum quantity demanded Q_min, and consumers' price sensitivity, k_P. If the data indicate bimodality, we also optimize over the generic fraction x.

We model heterogeneity among firms by adding random variables ζ_i and ξ_i to the parameters ΔQ_ad and Q_min, respectively. These random variables are drawn from a normal distribution with mean zero and respective standard deviations ϵ_1 and ϵ_2. We interpret ϵ_1 as the variation in the quality of advertising messaging and ϵ_2 as the variation in natural demand for the firms' products, and we also optimize their values.

We must choose starting "seeds" for the Nelder–Mead algorithm since it is a local optimization method. We do this by first extracting two modes from the price data, one which corresponds to the lower advertising investment group (generics) and the other the higher advertising investment group (name brands). We then choose seed parameters such that the model's predicted price distribution matches up with those two modes. For a more detailed description of the initialization of the algorithm, see the SM.

Figure 1 provides a few examples of fits for products that had more than 10,000 transactions. These examples also demonstrate the variety of products within the data set. We see there is qualitative agreement between the model's predicted price distributions and the empirical price distributions.

4.2. Statistics. We attempt to validate our model by fitting theoretical price distributions to empirical data provided by the Nielsen Corporation [7]. We use two tests, the Kolmogorov–Smirnov (KS) test and Hartigan's Dip Test, to assess the quality of our fits. See Figure 1 for a sample of model fits to data.

The KS test generates the probability that two samples come from the same underlying distribution by calculating the maximum absolute difference between their cumulative distribution functions (CDFs). Here, a large difference implies a low probability that the two data sets come from the same distribution. For a majority (58%) of our model fits to the top 500 products, we fail to reject the null hypothesis (samples from same underlying distribution) at a significance level of 0.05: the data and the model prediction may come from the same distribution.

Hartigan's Dip Test assesses whether a distribution is unimodal by comparing the CDF of the distribution to a unimodal test distribution [18]. A large difference between the distribution in question and the test distribution indicates a low probability of the distribution being unimodal. We apply Hartigan's Dip Test to the 500 products with the most entries in the database, and find that 46% have price distributions inconsistent with unimodality at a significance level 0.05. If price distributions are linked to advertising expenditures, as our model indicates, then almost half the products have a multimodal (bimodal or higher number of modes[8]) advertising distribution. For other products, unimodality could not be rejected, but data may not be inconsistent with bimodality. See the SM for the full distribution of p-values.

5. Discussion. The theory we present provides a possible explanation for the segmentation of commodity-product sectors into "name-brand" and "generic" products. We speculate that similar explanations might exist for other contexts where hierarchy emerges as a result of competition, or where interactions between individual agents can lead to global patterns [29]. For example, competition for a mate [24, 9, 20] and competition for resources [37, 19, 2] can both result in hierarchies observed in the natural world. Our model might be adapted to yield insight into such phenomena.

5.1. Limitations. In creating a highly simplified model, we have inevitably made some assumptions that limit its generality. These include the following assumptions.

- We assumed that advertising was persuasive and, hence, that quantity demanded increased uniformly across all price levels as advertising increased. In cases where advertising is informative, however, one would expect the slope of the demand curve to increase, instead of simply shifting vertically.
- We chose to leave the development of brand loyalty out of our model. This could presumably be captured through a demand curve that becomes more

[8]We suspect that an extension of this model to allow stronger within-segment competition (i.e., name brands compete more strongly with each other than with generics) would lead to additional modes.

inelastic as loyalty increases.
- We excluded spillover effects from "generic advertising," whereby advertising leads to increases in demand for all companies selling a similar product [28, 3]. We expect that this would increase profit for all companies but not affect the bimodal segmentation our current model predicts.
- We assumed advertising has a stable and lasting impact. Our model treats the benefits of advertising as arising instantaneously, an approximation that is only merited when the time scale of interest is much longer than the advertising's "half-life" in the consumer environment.
- We assumed the existence of many producers selling similar products. Some of our arguments would not be valid in the case of an oligopoly, where there are only a handful of producers.
- We approximated demand curves as linear, but of course these could (and likely do) take on more complex forms for real products.

In addition to the limitations of our modeling approach, the data set we examined also contains some biases that should be pointed out. Most saliently, the price distributions we examined are the result of different vendors selling identical products for different prices: this means that branding is really present at the vendor level, slightly different from the most direct and natural interpretation of the model. Also, a large fraction of entries in the database are food and other consumable products, since these are purchased more frequently than durable goods. Consumables might have a different market structure than products in nonfood markets (e.g., electronics, health care, housing, etc.).

5.2. Conclusions. We have presented a simple mathematical model for competition among firms on the basis of advertising. Despite the model's simplicity, a surprisingly robust prediction emerges: products split into "name-brand" and "generic" groups. This prediction appears to be largely consistent with data both in a qualitative sense (many products have nonunimodal price distributions) and a quantitative sense (theoretical price distributions from the model are consistent with empirical price distributions), even without a more detailed and accurate model.

Advertising has a large macroeconomic impact on corporate profits, market efficiency, and consumer welfare. The segmentation we report contrasts starkly with (often implicit) assumptions of smooth, singly peaked functions for economic metrics. We hope that our work helps refine intuition and inspires further inquiries into this intriguing aspect of free market dynamics.

Data Availability. The data that support the findings of this study are available from The Nielsen Company (US), LLC, but restrictions apply to the availability of these data, which were used under license for the current study and so are not publicly available. Data are however generally available for scientific research with an institutional or individual subscription [7].

Acknowledgments. We thank the Nielsen Corporation for their generosity in making their proprietary data available for scientific research. We also thank Vicky Chuqiao Yang for useful conversations.

Author Contributions. DMA and JDJ conceived of the research. JDJ performed the majority of the model analysis and numerical integration/simulation. AMR helped write code for numerical simulation. DMA and JDJ wrote the paper together.

Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts

Center for Marketing Data Center at The University of Chicago Booth School of Business.

The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

REFERENCES

[1] A. M. Abernethy and D. D. Butler, *Advertising information: Services versus products*, J. Retailing, 68 (1992), p. 398. (Cited on p. 626)

[2] P. Auger, R. B. De La Parra, and E. Sánchez, *Hawk-dove game and competition dynamics*, Math. Comput. Model., 27 (1998), pp. 89–98. (Cited on p. 636)

[3] F. M. Bass, A. Krishnamoorthy, A. Prasad, and S. P. Sethi, *Generic and brand advertising strategies in a dynamic duopoly*, Marketing Sci., 24 (2005), pp. 556–568. (Cited on p. 637)

[4] G. S. Becker and K. M. Murphy, *A simple theory of advertising as a good or bad*, Quart. J. Econom., 108 (1993), pp. 941–964. (Cited on p. 625)

[5] F. Bloch and D. Manceau, *Persuasive advertising in Hotelling's model of product differentiation*, Internat. J. Indust. Organiz., 17 (1999), pp. 557–574. (Cited on p. 625)

[6] D. Braithwaite, *The economic effects of advertisement*, Econom. J., 38 (1928), pp. 16–37. (Cited on p. 626)

[7] Chicago Booth Kilts Center for Marketing, *Nielsen Consumer Panel Data*, 2014, https://www.chicagobooth.edu/research/kilts/datasets/nielsen; accessed October 25, 2016 (registration required). (Cited on pp. 627, 634, 636, 637)

[8] A. Clauset, C. R. Shalizi, and M. E. J. Newman, *Power-law distributions in empirical data*, SIAM Rev., 51 (2009), pp. 661–703, https://doi.org/10.1137/070710111. (Cited on p. 628)

[9] S. M. Clifton, R. I. Braun, and D. M. Abrams, *Handicap principle implies emergence of dimorphic ornaments*, Proc. Roy. Soc. B Biol. Sci., 283 (2016), art. 20161970. (Cited on pp. 633, 636)

[10] B. Crettez, N. Hayek, and G. Zaccour, *Existence and uniqueness of optimal dynamic pricing and advertising controls without concavity*, Oper. Res. Lett., 46 (2018), pp. 199–204. (Cited on p. 626)

[11] R. Crisp, *Persuasive advertising, autonomy, and the creation of desire*, J. Business Ethics, 6 (1987), pp. 413–418. (Cited on p. 625)

[12] A. K. Dixit and V. Norman, *Advertising and welfare*, Bell J. Econom., 9 (1976), pp. 1–17. (Cited on p. 625)

[13] G. M. Erickson, *A differential game model of the marketing-operations interface*, European J. Oper. Res., 211 (2011), pp. 394–402. (Cited on p. 626)

[14] L. Esteban, A. Gil, and J. M. Hernandez, *Informative advertising and optimal targeting in a monopoly*, J. Indust. Econom., 49 (2001), pp. 161–180. (Cited on p. 626)

[15] C. Fershtman, *Goodwill and market shares in oligopoly*, Economica, 51 (1984), pp. 271–281. (Cited on p. 626)

[16] J. W. Friedman, *Advertising and oligopolistic equilibrium*, Bell J. Econom., 14 (1983), pp. 464–473. (Cited on p. 626)

[17] L. Gori, M. Sodini, and L. Fanti, *A nonlinear Cournot duopoly with advertising*, Chaos Solitons Fractals, 79 (2015), pp. 178–190. (Cited on p. 626)

[18] J. A. Hartigan and P. Hartigan, *The dip test of unimodality*, Ann. Statist., 13 (1985), pp. 70–84. (Cited on p. 636)

[19] M. E. Hibbing, C. Fuqua, M. R. Parsek, and S. B. Peterson, *Bacterial competition: Surviving and thriving in the microbial jungle*, Nature Rev. Microbiol., 8 (2010), pp. 15–25. (Cited on p. 636)

[20] M. Kirkpatrick, *Sexual selection and the evolution of female choice*, Evolution, 36 (1982), pp. 1–12. (Cited on p. 636)

[21] N. Lindstädt and O. Budzinski, *Newspaper vs. Online Advertising: Is There a Niche for Newspapers in Modern Advertising Markets?*, IME Working Paper 113, University of Southern Denmark, Department of Environmental and Business Economics (IME), Esbjerg, 2011. (Cited on p. 626)

[22] C. Marinelli and S. Savin, *Optimal distributed dynamic advertising*, J. Optim. Theory Appl., 137 (2008), pp. 569–591. (Cited on p. 626)

[23] D. Marković and C. Gros, *Power laws and self-organized criticality in theory and nature*, Phys. Rep., 536 (2014), pp. 41–74. (Cited on p. 628)

[24] E. L. McCullough, C. W. Miller, and D. J. Emlen, *Why sexually selected weapons are not*

ornaments, Trends Ecology Evol., 31 (2016), pp. 742–751. (Cited on p. 636)

[25] J. A. Nelder and R. Mead, *A simplex method for function minimization*, Computer J., 7 (1965), pp. 308–313. (Cited on p. 635)

[26] P. Nelson, *The economic consequences of advertising*, J. Business, 48 (1975), pp. 213–241. (Cited on p. 625)

[27] L. M. Nichols, *Advertising and economic welfare*, Amer. Econom. Rev., 75 (1985), pp. 213–218. (Cited on p. 625)

[28] G. Norman, L. Pepall, and D. Richards, *Generic product advertising, spillovers, and market concentration*, Amer. J. Agricultural Econom., 90 (2008), pp. 719–732. (Cited on p. 637)

[29] S. E. Page, *Aggregation in agent-based models of economies*, Knowledge Engrg. Rev., 27 (2012), pp. 151–162. (Cited on p. 636)

[30] R. Schmalensee, M. Armstrong, R. D. Willig, and R. H. Porter, *Handbook of Industrial Organization*, Vol. 3, Elsevier, 1989. (Cited on p. 625)

[31] S. P. Sethi, *Deterministic and stochastic optimization of a dynamic advertising model*, Optimal Control Appl. Methods, 4 (1983), pp. 179–184. (Cited on p. 626)

[32] H. Shaban, *Digital advertising to surpass print and TV for the first time, report says*, The Washington Post, February 20, 2019, https://www.washingtonpost.com/technology/2019/02/20/digital-advertising-surpass-print-tv-first-time-report-says/. (Cited on p. 625)

[33] H. Simon, *ADPULS: An advertising model with wearout and pulsation*, J. Marketing Res., 19 (1982), pp. 352–363. (Cited on p. 626)

[34] G. J. Stigler, *The economics of information*, J. Political Econ., 69 (1961), pp. 213–225. (Cited on p. 625)

[35] G. J. Stigler and G. S. Becker, *De gustibus non est disputandum*, Amer. Econom. Rev., 67 (1977), pp. 76–90. (Cited on p. 625)

[36] L. G. Telser, *Advertising and competition*, J. Political Econ., 72 (1964), pp. 537–562. (Cited on p. 625)

[37] D. Tilman, *Resource competition between plankton algae: An experimental and theoretical approach*, Ecology, 58 (1977), pp. 338–348. (Cited on p. 636)

[38] M. Tribe and R. Alpine, *Scale economies and the "0.6 rule,"* Engrg. Costs Prod. Econom., 10 (1986), pp. 271–278. (Cited on p. 628)