Advertising and Collusion in Retails Markets

Gea Myoung Lee, Kyle Bagwell
March 2008

Advertising and Collusion in Retail Markets

Kyle Bagwell and Gea M. Lee

March 29, 2008
Preliminary Draft

Abstract

We consider non-price advertising by retail firms that are privately informed as to their respective production costs. We first analyze a static model. We construct an advertising equilibrium, in which informed consumers use an advertising search rule whereby they buy from the highest-advertising firm. Consumers are rational in using the advertising search rule, since the lowest-cost firm advertises the most and also selects the lowest price. Even though the advertising equilibrium facilitates productive efficiency, we establish conditions under which firms enjoy higher expected profit when advertising is banned. Consumer welfare falls in this case, however. We next analyze a dynamic model in which privately informed firms interact repeatedly. In this setting, firms may achieve a collusive equilibrium in which they limit the use of advertising, and we establish conditions under which optimal collusion entails pooling at zero advertising. More generally, full or partial pooling is observed in optimal collusion. In summary, non-price advertising can promote product efficiency and raise consumer welfare; however, firms often have incentive to diminish advertising competition, whether through regulatory restrictions or collusion.

1 Introduction

Modern theoretical analyses of collusion emphasize collusion in prices or quantities. This emphasis is appropriate for many applications; however, collusion may also occur with respect to instruments of non-price competition. One possibility of particular interest is that firms select their advertising levels in a collusive fashion. This possibility has not received significant theoretical attention.[1]

One reason may be that the empirical literature on collusion and advertising offers somewhat mixed findings.[2] Ferguson (1974) argues that advertising activity is publicly observable and thus that collusion in advertising is feasible; and Cable (1972), Greer (1971) and Sutton (1974) emphasize the possibility of collusion in advertising among firms in highly concentrated markets, in their interpretations of the empirical relationship between advertising and concentration. Simon (1970) and Scherer (1980), however, argue that advertising activities are difficult to assess and monitor, and thus suggest that collusion in advertising may be difficult to achieve. More recently, Gasmi, Laffont and Vuong (1992) argue that Coca-Cola and Pepsi-Cola colluded in advertising and possibly price over a sample period that covers the late 1970s and early 1980s, and Kadiyali (1996) reports evidence that Kodak and Fuji colluded in price and advertising in the U.S. photographic film industry in the 1980s. But Symeonidis (2000) reports an absence of collusion in non-price variables like advertising in his study of U.K. manufacturing cartels.

In the specific context of retail markets, however, some interesting empirical relationships between advertising and prices have been identified. The classic study is by Benham (1972). Examining the retail eyeglass industry in the U.S. in the 1960s, he reports that retail prices were higher in states that prohibited all advertising than in states that had no restrictions on advertising; moreover, prices were only slightly higher in states that allowed just non-price advertising than in states that also allowed price advertising. Evidently, the ability to advertise, even if only in a non-price form, results in lower prices. Cady (1976) documents similar relationships in the U.S. retail market for prescription drugs in 1970. At a broad level, this work suggests that retail firms may gain if they are able to limit advertising. In the absence of a state law that prohibits advertising, retail firms thus have incentive to achieve a collusive agreement in which combative advertising is reduced.

Bagwell and Ramey (1994a) develop a complete-information model of retail competition with which to interpret Benham's findings. In their model, some consumers can identify the highest-advertising firm, while other consumers do not observe advertising levels. The former (latter) consumers are referred to as informed (uninformed) consumers. Each consumer possesses a downward-sloping demand function and lacks direct information about firms' prices: a consumer observes a firm's price only after choosing to visit that firm. Bagwell and Ramey compare two equilibria. In a random equilibrium, consumers ignore advertising and choose firms at random. Firms do not ad-

---

[1] For exceptions, see Friedman (1983) and Stigler (1968). Friedman characterizes open-loop Nash equilibria in a repeated game of advertising and quantity competition, while Stigler compares cartels that collude in advertising and compete in price with those that collude in price and compete in advertising. See also Nocke (2007) for a recent analysis of collusive equilibria in a dynamic game of investment, where investment may be thought of as quality-improving R&D or persuasive advertising.

[2] For a comprehensive survey of the economic analysis of advertising, see Bagwell (2007).

vertise and enjoy symmetric market shares. By contrast, in an advertising equilibrium, the informed consumers go to the firm that advertises the most. Firms then use a symmetric mixed strategy, in which higher advertising choices are paired with greater investments in cost reduction and thus lower prices. Informed consumers are then rational in visiting the highest-advertising retailer, since this retailer also offers the lower price. Bagwell and Ramey include an initial entry stage and show that, in the advertising equilibrium, the market is more concentrated, prices are lower, and social welfare is higher. If the random equilibrium is associated with a setting in which advertising is banned, these findings are broadly consistent with the empirical patterns that Benham reports.

In this paper, we modify the Bagwell-Ramey model in two key respects. Our first modification is to "purify" the model and assume that each firm has private information about its costs of production. Specifically, we consider a model with a continuum of possible cost types, where cost types are iid across firms. In the corresponding static model, we characterize an advertising equilibrium in which firms use pure strategies and lower-cost firms advertise strictly more than do higher-cost firms. The advertising equilibrium again may be compared with the random equilibrium in which no firm advertises. Our second modification is to allow that firms interact repeatedly over an infinite horizon, where each firm's cost type is iid over time. With this second modification, we may consider any self-enforcing collusive agreement among firms. Thus, in our modified model, the search for an optimal collusive equilibrium among firms entails significantly more than a comparison of the random and advertising equilibria.

In our analysis of the static model, we capture two notions of purification. First, in the special case in which the support of possible cost types is small, we report that the distribution of advertising levels in the pure-strategy advertising equilibrium of the incomplete-information game is approximately the same as the distribution of advertising levels in the mixed-strategy advertising equilibrium of the complete-information game.[3] Correspondingly, we show that the main findings of Bagwell and Ramey directly extend to the private-information setting, if the support of possible costs is sufficiently small. Second, in the general case in which the support of possible cost types may be large, we establish conditions under which the main findings derived in the complete-information game arise also in the incomplete-information game. This second notion of purification sometimes requires additional structure on the distribution and demand functions.

We develop our results for the static game while allowing for a general support of possible cost types. As mentioned, we establish that an advertising equilibrium exists, in which lower-cost firms advertise more and price lower than do higher-cost firms. We then establish three further results. First, for any given number of firms, if the distribution of types is log-concave and demand is sufficiently inelastic, then firms earn higher expected profit in the random than in the advertising equilibrium. The second result follows directly from the first: when the number of firms is endogenous, if the distribution of types is log-concave and demand is sufficiently inelastic, more firms enter when the random equilibrium is anticipated. Finally, without making any assumption

---

[3]Bagwell and Ramey (1994a) report a similar finding when the cost of advertising is private information and varies slightly across firms.

on the distribution of types or the elasticity of demand, we show that social surplus is weakly higher in the advertising than in the random equilibrium, when the number of firms is endogenous. In fact, social surplus is strictly higher in the advertising equilibrium if at least two firms enter in that equilibrium. We thus establish a general sense in which Bagwell and Ramey's main findings extend to the private-information setting. We emphasize, however, that the first two results mentioned now employ additional assumptions on the distribution of types and the elasticity of demand.

We also compare the advertising equilibrium with another benchmark. In particular, we follow Varian (1980) and suppose that informed consumers observe prices and buy from the lowest-priced firm while uninformed consumers pick a firm at random. Following Spulber (1995) and Bagwell and Wolinsky (2002), we modify Varian's model and allow that firms are privately informed about their production costs.[4] Let us refer to the (symmetric) equilibrium of this game as the pricing equilibrium. For any fixed number of firms, we show that firms earn higher expected profit in the pricing equilibrium than in the advertising equilibrium. This is perhaps surprising, since competition in advertising is sometimes argued to be less aggressive than competition in prices. As we discuss, the key intuition is that price competition induces greater in-store demand from consumers and thus elevates the size of expected information rents for firms.

With an analysis of the benchmark model in place, we are able to offer a more complete comparison across different advertising regulatory regimes. Provided that the market always has at least two firms, our results indicate that the average transaction price is lowest in the pricing equilibrium, somewhat higher in the advertising equilibrium, and higher yet in the random equilibrium. Likewise, when the number of firms is endogenous, social welfare is highest in the pricing equilibrium, somewhat lower in the advertising equilibrium, and lower yet in the random equilibrium. If we associate the pricing equilibrium with a setting in which price advertising is allowed, the advertising equilibrium with a setting in which only non-price advertising is allowed, and the random equilibrium with a setting in which all advertising is banned, then our results are broadly consistent with Benham's findings.

We next examine the comparative-statics properties of the advertising equilibrium. When the number of informed consumers is increased, advertising increases for all types other than the highest type. Intuitively, firms advertise more heavily when the "prize" from advertising the most is increased. Interestingly, the effect on advertising of an increase in the number of firms depends on a firm's cost type: lower-cost firms compete more aggressively and increase their advertising, but higher-cost firms perceive a reduced chance of winning the informed consumers and advertise less. An implication is that the support of observed advertising levels may be larger in markets with a greater number of firms. We also find that, for all types other than the lowest type, if the number of firms is sufficiently large, the equilibrium level of advertising is negligible. Finally, building on Hopkins and Korneinko's (2007) analysis of all-pay auctions, we show that, if the cost distribution shifts to make lower-cost types more likely in the sense of the monotone likelihood ratio order, then

---

[4]Bagwell and Wolinsky follow Varian and assume that each consumer possesses an inelastic demand function. We generalize this analysis slightly and allow for downward-sloping demand functions. Spulber considers a related model in which all consumers are informed.

lower-cost firms advertise more while higher-cost firms become discouraged and advertise less.[5]

We then turn to consider advertising in the associated repeated game. Assuming that informed consumers go to the highest-advertising firm within any given period and that advertising selections are publicly observed by firms, we focus on the symmetric perfect public equilibria (SPPE) of our repeated game with private information. For this class of equilibria, our goal is to characterize the optimal form of collusion in advertising among a fixed number of firms.[6] We note that SPPE include a wide range of equilibrium behaviors. Firms may repeatedly play the (non-cooperative) advertising equilibrium of the static game, and patient firms may also enforce zero advertising in all periods. In the latter case, collusion among firms is used to implement repeatedly the random equilibrium. The random equilibrium is then achieved as a self-enforcing ban on advertising rather than as a consequence of a legal ban on advertising. Patient firms may also implement other stationary advertising strategies, including advertising schedules that take the form of step functions. A further possibility is that firms implement an SPPE that entails non-stationary play, with firms moving between cooperative and war phases in their advertising conduct.

When firms collude in private-information settings, two kinds of incentive constraints arise.[7] First, each firm must not gain by undertaking an "on-schedule deviation," whereby a firm with one cost type deviates and mimics the behavior that is prescribed for this firm when it has a different cost type. The on-schedule incentive constraint is analogous to the standard truth-telling constraint encountered in mechanism-design problems. An important feature of an on-schedule deviation is that no other firm would be aware that a deviation actually occurred, since other firms would infer that the firm drew the cost type for which the observed behavior is prescribed in equilibrium. The second kind of deviation is called an "off-schedule deviation." An off-schedule deviation occurs when a firm takes an action that is not specified in equilibrium for any of its possible cost types. Importantly, an off-schedule deviation is publicly observed as a deviation. As in standard repeated games, an off-schedule deviation is punished harshly; thus, sufficiently patient firms will not undertake off-schedule deviations.

Colluding firms face interesting trade-offs when selecting an optimal collusive scheme. Suppose firms contemplate the repeated use of the advertising equilibrium of the static game. An advantage of this scheme is that it maximizes productive efficiency: in each period, lower-cost firms advertise at strictly higher levels, and so the informed consumers are allocated to the lowest-cost firm. A disadvantage of this scheme, however, is that firms' profits are reduced by high advertising expenditures. Firms may thus look for some way to keep the productive-efficiency advantage while reducing advertising expenditures. They might thus consider a strictly decreasing advertising schedule that is "flatter" and involves lower levels of advertising. Such a schedule, however, will induce higher-cost types to raise their advertising and mimic lower-cost types, unless higher advertising selections

---

[5]This finding contrasts interestingly with the standard monotone comparative statics result for first-price auctions. See, for example, Athey (2002) and Lebrun (1998).

[6]In the stage game, sequential search is not allowed, and firms are thus able to select their respective monopoly prices. We therefore embed monopoly pricing into the profit functions and focus on collusion in advertising.

[7]The discussion here follows Athey, Bagwell and Sanchirico (2004) and Athey and Bagwell (2001).

result in some future cost. Given our focus on SPPE, any future cost must be experienced symmetrically by all firms. The future cost may thus take the form of a future advertising "war" in which higher and less profitable advertising schedules are employed. This discussion points to two general themes. First, there is a substitutability between current-period advertising and future advertising wars. Second, the productive-efficiency benefits that are associated with sorting can be enjoyed only if the informational cost of high current or future advertising levels is also experienced.

Our formal analysis builds on these themes. We show that an optimal SPPE always exists that is stationary (i.e., that does not use wars). This result holds for all demand and distribution functions. It thus confirms at a general level that future advertising wars are a redundant instrument. We also characterize an optimal SPPE that is stationary. In particular, if the distribution function is log-concave and the demand function is sufficiently inelastic, then the optimal SPPE for sufficiently patient firms entails pooling at zero advertising for all cost types in all periods. Thus, while the repeated game allows for a wide range of SPPE advertising behaviors, under some conditions, the optimal SPPE is a self-enforcing agreement among firms to eliminate combative advertising.

We emphasize that this result requires patient firms and assumes sufficiently inelastic demand. Firms must be patient in order to resist undertaking an off-schedule deviation and advertising a positive amount. For patient firms, the immediate gain in profit would be overwhelmed by the loss in future profit that would ensue. For example, such a deviation might trigger reversion to the advertising equilibrium of the static game in all future periods. Likewise, for other demand functions, the optimal SPPE may not entail zero advertising by all types. We show, though, that under general conditions the optimal SPPE has partial rigidity (i.e., intervals of cost types with pooling). Further, for any demand function, in the special case where the support of possible cost types is sufficiently small, the optimal SPPE entails pooling at zero advertising by all types. Finally, we also consider the case of a uniform distribution of types and a CES demand function. For this case, we show that the optimal SPPE again entails pooling at zero advertising by all types, if the elasticity of demand does not exceed a critical level where the critical level is higher when the support of possible cost types is smaller.

Our analysis of the repeated advertising game is closely related to work by Athey, Bagwell and Sanchirico (2004).[8] They consider a repeated game in which firms have private cost shocks and collude in pricing. When the distribution of cost types is log-concave, if demand is sufficiently inelastic, the optimal SPPE for sufficiently patient firms is a stationary equilibrium in which firms always select the same price, regardless of their respective cost types. We find a similar force in favor of pooling when firms collude in advertising. Athey, Bagwell and Sanchirico also establish that an optimal SPPE exists that is stationary, if demand is sufficiently inelastic. In our model of collusion in advertising, for general demand functions, an optimal SPPE exists that is stationary. The game considered by Athey. Bagwell and Sanchirico may be thought of as a repeated first-price

---

[8]See also McAfee and McMillan (1992) for a related theory of identical bidding among collusive bidders. They develop their results for a first-price auction in a static model. Our model of advertising is analogous to an all-pay auction, and we also present a dynamic analysis. For other analyses of repeated games with private information in which SPPE are analyzed, see Bagwell and Staiger (2005), Hanazono and Yang (2007) and Lee (2007).

(procurement) auction, while the repeated advertising game that we analyze here is analogous to a repeated all-pay auction.

We next return to the static model and extend our analysis to allow for sequential search. If demand is sufficiently inelastic or if the cost of sequential search is sufficiently high, then our results are maintained without modification. If these conditions do not hold, however, then higher-cost firms must "limit price" (i.e., price below their monopoly prices), in order to deter sequential search.[9] An advertising equilibrium then continues to exist, if the support of possible cost types is not too large and the number of informed consumers is not too great. In this equilibrium, informed consumers use observed advertising behavior to locate the lowest price, and limit pricing by higher-cost firms ensures that uninformed consumers do not gain from actually undertaking sequential search. We argue as well that the possibility of sequential search may even strengthen our results, by raising the relative profitability of the random equilibrium.

We conclude by briefly discussing two other extensions. First, while advertising entails money burning in our main analysis, it is also plausible that advertising may directly enter the demand function. Second, our analysis of the repeated game focuses on SPPE, and we briefly discuss some potential issues that might arise in an analysis of asymmetric PPE.

The paper is organized as follows. Section 2 contains the analysis of the static game. The repeated game is examined in Section 3. Optimal collusion is characterized in Section 4. In Section 5, we extend the static model to allow for sequential search. Section 6 contains a brief discussion of other extensions, and Section 7 concludes. Remaining proofs are in the Appendix.

2 The Static Game

In this section, we define a static game in which a fixed number of firms compete through advertising for market share. Firms are privately informed as to their respective costs, and each firm's advertising choice may signal its costs, and thus its price, to those consumers who are informed of advertising activities. We establish the existence of an advertising equilibrium, in which informed consumers visit the firm with the highest level of advertising. We compare the expected profit earned by firms in the advertising equilibrium with that which they earn in a random equilibrium, wherein all consumers pick firms at random. We then endogenize the number of firms and compare market concentration and social welfare across the two equilibria. Next, we compare the advertising equilibrium with the pricing equilibrium of a benchmark model, in which some consumers observe all prices. Finally, we provide a comparative-statics analysis of the advertising equilibrium.

2.1 The Model

We assume that $N \geq 2$ ex ante identical firms compete for sales in a homogeneous-good market. Each firm i is privately informed of its unit cost level θ_i. Cost levels are iid across firms, and cost type

---

[9]Our analysis here builds on Reinganum (1979) and Bagwell and Ramey (1996). Reinganum examines sequential search in a model with privately informed firms that are not allowed to advertise. Bagwell and Ramey examine sequential search when advertising is allowed but private information is absent.

θ_i is drawn from the support $[\underline{\theta}, \overline{\theta}]$ according to the twice-continuously differentiable distribution function, $F(\theta)$, where $\overline{\theta} > \underline{\theta} \geq 0$. The density $f(\theta) \equiv F'(\theta)$ is positive on $[\underline{\theta}, \overline{\theta}]$. After firms learn their respective cost types, they simultaneously choose their levels of advertising. A pure strategy for firm i is a function, $A_i(\theta_i)$, that maps from the set of cost types $[\underline{\theta}, \overline{\theta}]$ to the set of possible advertising expenditures $\mathbb{R}_+ \equiv [0, \infty)$. For simplicity, we assume A_i is continuously differentiable except at perhaps a finite number of points where the function jumps. Given a strategy profile $[A_1, ..., A_N]$, let A_{-i} denote the strategies of firms other than i and let $A_{-i}(\theta_{-i})$ denote the vector of these firms' selections when their cost types are given by the $(N-1)$-tuple θ_{-i}.

There is a unit mass of consumers, where each consumer possesses a twice-continuously differentiable demand function $D(p)$ that satisfies $D(p) > 0 > D'(p)$ over the relevant range of prices p. Following Bagwell and Ramey (1994a), we assume that advertising is a dissipative expense that does not directly affect demand. Consumers cannot observe prices prior to their visitation decision; thus, prices cannot be directly communicated in the market. Consumers are divided into two groups. A fraction I of consumers observe firms' advertising expenses.[10] Given this information, informed consumers form beliefs as to firms' cost types and determine a search (visitation) strategy. For example, informed consumers may use an *advertising search rule*, whereby a consumer goes to the firm that advertises the most. The remaining fraction $U = 1 - I$ are uninformed. Uninformed consumers do not observe advertising expenditures and always follow a *random search rule*, whereby a consumer randomly chooses which firm to visit.

The interaction between firms and consumers is represented by the following static game: (i) firms learn their own cost types, (ii) firms make simultaneous choices of advertising and price, and (iii) given any advertising information, each consumer chooses a firm to visit, observes that firm's price and makes desired purchases given this price. Note that a consumer is assumed to visit only one firm.[11] This simplifies our analysis, since it ensures that each firm chooses the monopoly price that is associated with its cost type. Consequently, we assume that monopoly prices are selected and focus on advertising selections.

We next describe a firm's expected profit. A firm's net revenue is $r(p, \theta) \equiv (p - \theta)D(p)$ (excluding advertising expense) when it has cost type θ, sets the price p and captures the entire unit mass of consumers. We assume that $r(p, \theta)$ is strictly concave in p with a unique maximizer $p(\theta) = \arg\max_p r(p, \theta)$. It follows that the monopoly price $p(\theta)$ strictly increases in θ whereas $r(p(\theta), \theta)$ strictly decreases in θ. We further assume that the price at the top has a positive margin: $p(\overline{\theta}) > \overline{\theta}$. The market share for firm i, denoted by m_i, maps from $\mathbb{R}_+^N$ to $[0, 1]$. Given the search rule used by informed consumers, m_i is determined by the vector of advertising levels selected by firm i and its rivals. If firm i has cost type θ_i, its interim-stage market share is $E_{\theta_{-i}}[m_i(A_i(\theta_i), A_{-i}(\theta_{-i}))]$. Embedding the monopoly price $p(\theta_i)$ into the revenue function, we may define the interim-stage net revenue for firm i by $R(A_i(\theta_i), \theta_i; A_{-i}) \equiv r(p(\theta_i), \theta_i)E_{\theta_{-i}}[m_i(A_i(\theta_i), A_{-i}(\theta_{-i}))]$. Firm i's expected revenue is $E_{\theta_i}R(A_i(\theta_i), \theta_i; A_{-i})$, and firm i's expected profit is thus $E_{\theta_i}[R(A_i(\theta_i), \theta_i; A_{-i}) - A_i(\theta_i)]$.

---

[10] We assume that informed consumers observe advertising levels for simplicity. In fact, all of our results hold under the assumption that informed consumers observe only the identity of the highest-advertising firm(s).

[11] We extend the analysis to allow for sequential search in Section 5.

For the static game, we are interested in Symmetric Perfect Bayesian Equilibria in which firms select their monopoly prices and uninformed consumers use the random search rule. We thus define an *equilibrium* as a profile $[A_1, ... A_N]$ and a belief function and search rules for consumers that collectively satisfy four conditions. First, given the market share function, m_i, that is induced by consumers' search rules, the profile $[A_1, ... A_N]$ is such that, for all i and θ_i, $A_i(\theta_i) \in \arg\max_{a_i} [R(a_i, \theta_i; A_{-i}) - a_i]$. Second, given an observed advertising level a_i by firm i, informed consumers use Bayes' Rule whenever possible (i.e., whenever $a_i = A_i(\theta_i)$ for some $\theta_i \in [\underline{\theta}, \overline{\theta}]$) in forming their beliefs as to firm i's cost type θ_i and thus price $p(\theta_i)$. Third, for any observed profile of advertising levels $[a_1, ..., a_N]$, given their beliefs, the informed consumers' search rule directs them to the firm or firms with the lowest expected price. Finally, firms' advertising strategies are symmetric: $A_i = A$ for all i.

Given symmetry, we can simplify notation somewhat. We can now define firm i's interim-stage market share as $M(A(\theta_i); A) \equiv E_{\theta_{-i}} [m_i(A_i(\theta_i), A_{-i}(\theta_{-i}))]$. Similarly, we can define firm i's interim-stage profit and net revenue as follows:

$$
\begin{aligned}
\Pi(A(\theta_i), \theta_i; A) &\equiv r(p(\theta_i), \theta_i) M(A(\theta_i); A) - A(\theta_i). \\
&\equiv R(A(\theta_i), \theta_i; A) - A(\theta_i).
\end{aligned}
$$

We note that the interim-stage profit function satisfies a single-crossing property: higher types are less willing to engage in higher advertising to increase expected market share.[12] For here and later use, we now write interim-stage profit in direct-form notation, ignoring subscript i: if a firm of type θ picks an advertising level $A(\widehat{\theta})$ when its rivals employ the strategy A, then we define $\Pi(\widehat{\theta}, \theta; A) \equiv \Pi(A(\widehat{\theta}), \theta; A)$, $M(\widehat{\theta}; A) \equiv M(A(\widehat{\theta}); A)$ and $R(\widehat{\theta}, \theta; A) \equiv R(A(\widehat{\theta}), \theta; A)$.

We are primarily interested in two kinds of equilibria. In an *advertising equilibrium*, informed consumers use the advertising search rule, whereby they go to the firm that advertises the most.[13] Since $p(\theta)$ is strictly increasing, such equilibria can exist only if the advertising schedule A is nonincreasing, so that higher-advertising firms have lower costs and thus offer lower prices. In a *random equilibrium*, informed consumers ignore advertising and use the random search rule. A random equilibrium thus can exist only if firms maximize expected profits and do not advertise (i.e., $A \equiv 0$). We explore these equilibria in the next two subsections.

2.2 Advertising Equilibrium

In an advertising equilibrium, informed consumers use the advertising search rule while uninformed consumers are randomly distributed across all N firms. We now report the following existence and uniqueness result.

---

[12] When a firm increases its advertising level, it may confront a trade off between the larger advertising expense, a_i, and the consequent higher expected market share, $M(a_i; A)$. When the interim-stage profit is held constant, the slope $da_i/dM(a_i; A)$ is given by $r(p(\theta_i), \theta_i)$, which is strictly decreasing in θ_i.

[13] Under the advertising search rule, if several firms tie for the highest advertising level, then the informed consumers divide up evenly over those firms.

Proposition 1. *There exists a unique advertising equilibrium, and in this equilibrium $A(\theta)$ is strictly decreasing and differentiable and satisfies $A(\overline{\theta}) = 0$.*

Proof. We first derive the necessary features of an advertising equilibrium. Consider any $\widehat{\theta} < \theta$. The following incentive constraints are necessary:

$$r(p(\widehat{\theta}), \widehat{\theta})M(\widehat{\theta}; A) - A(\widehat{\theta}) \geq r(p(\widehat{\theta}), \widehat{\theta})M(\theta; A) - A(\theta)$$
$$r(p(\theta), \theta)M(\theta; A) - A(\theta) \geq r(p(\theta), \theta)M(\widehat{\theta}; A) - A(\widehat{\theta}).$$

Adding yields $[r(p(\widehat{\theta}), \widehat{\theta}) - r(p(\theta), \theta)][M(\widehat{\theta}; A) - M(\theta; A)] \geq 0$. Since $r(p(\theta), \theta)$ is strictly decreasing in θ, it is thus necessary that $M(\theta; A)$ is nonincreasing. It thus follows that $A(\theta)$ is nonincreasing. Further, given the advertising search rule, it is clear that $A(\theta)$ cannot be constant over any interval of types: by increasing its advertising an infinitesimal amount, a firm with a type on this interval would experience a discrete gain in its expected market share. Thus, $A(\theta)$ must be strictly decreasing, and consequently it is necessary that $M(x; A) = \frac{U}{N} + [1 - F(x)]^{N-1}I$. It thus follows that $M(\overline{\theta}; A) = \frac{U}{N}$. A firm with type $\overline{\theta}$ thus cannot be deterred from selecting zero advertising, and hence $A(\overline{\theta}) = 0$ is also necessary.

We next establish that $A(\theta)$ must be differentiable, and we also derive the necessary expression for $A'(\theta)$. Consider any $\widehat{\theta} < \theta$. Rearranging the incentive constraints presented above, we find that

$$\frac{r(p(\theta), \theta)[M(\theta; A) - M(\widehat{\theta}; A)]}{\theta - \widehat{\theta}} \geq \frac{A(\theta) - A(\widehat{\theta})}{\theta - \widehat{\theta}} \geq \frac{r(p(\widehat{\theta}), \widehat{\theta})[M(\theta; A) - M(\widehat{\theta}; A)]}{\theta - \widehat{\theta}}.$$

Taking limits as $\theta \to \widehat{\theta}$, and using the differentiability of $M(\theta; A) = \frac{U}{N} + [1 - F(x)]^{N-1}I$, we conclude that

$$A'(\theta) = r(p(\theta), \theta)\frac{\partial M(\theta; A)}{\partial \theta}.$$

When combined with the boundary condition $A(\overline{\theta}) = 0$, this differential equation may be solved to yield

$$A(\theta) = -\int_{\theta}^{\overline{\theta}} r(p(x), x)[\partial M(x; A)/\partial x]dx,$$

where $\frac{\partial M(x; A)}{\partial x} = -(N-1)[1 - F(x)]^{N-2}f(x)I < 0$ for all $x < \overline{\theta}$.

We now integrate by parts and establish that $A(\theta)$ must take the following unique form:

$$A(\theta) = R(\theta, \theta; A) - R(\overline{\theta}, \overline{\theta}; A) - \int_{\theta}^{\overline{\theta}} D(p(x))\left[\frac{U}{N} + [1 - F(x)]^{N-1}I\right]dx, \tag{1}$$

where $R(\overline{\theta}, \overline{\theta}; A) = r(p(\overline{\theta}), \overline{\theta})\frac{U}{N}$. Rearranging, we note that interim-stage profit for type θ then must be given as

$$\Pi(\theta, \theta; A) = R(\overline{\theta}, \overline{\theta}; A) + \int_{\theta}^{\overline{\theta}} D(p(x))\left[\frac{U}{N} + [1 - F(x)]^{N-1}I\right]dx. \tag{2}$$

Observe that interim-stage profit is positive for all $\theta \in [\underline{\theta}, \overline{\theta}]$.

The second step in our proof is to construct an advertising equilibrium using the $A(\theta)$ function defined in (1). Observe that $\Pi_1(\theta, \theta; A) = r(p(\theta), \theta)\frac{\partial M(\theta; A)}{\partial \theta} - A'(\theta) = 0$ when this function is used. It follows that no type θ will deviate by mimicking some other type $\widehat{\theta}$, since for all $\widehat{\theta} < \theta$ we have

$$
\begin{aligned}
\Pi(\theta, \theta; A) - \Pi(\widehat{\theta}, \theta; A) &= \int_{\widehat{\theta}}^{\theta} \Pi_1(x, \theta; A)dx \\
&= \int_{\widehat{\theta}}^{\theta} \left[\Pi_1(x, \theta; A) - \Pi_1(x, x; A)\right] dx \\
&= \int_{\widehat{\theta}}^{\theta} \int_{x}^{\theta} \Pi_{12}(x, y; A)dydx > 0,
\end{aligned}
$$

where the inequality follows from $\Pi_{12}(x, y; A) = D(p(y))(N-1)[1 - F(x)]^{N-2}f(x)I > 0$ for all $x < \overline{\theta}$. A similar argument ensures that $\Pi(\theta, \theta; A) > \Pi(\widehat{\theta}, \theta; A)$ for all $\widehat{\theta} > \theta$. Next, if no type $\theta > \underline{\theta}$ gains from deviating to $A(\underline{\theta})$, then a deviation to any $A > A(\underline{\theta})$ is also unattractive. Finally, since $A'(\theta) < 0$, the advertising search rule is optimal for informed consumers. ∎

Proposition 1 thus establishes the existence and uniqueness of an advertising equilibrium.[14] The advertising equilibrium acts as a fully sorting mechanism: firms truthfully reveal their cost types along the downward-sloping advertising schedule. The informed consumers behave rationally in the advertising model: the lowest-cost firm advertises the most and offers the lowest price, and the informed consumers purchase from the highest-advertising firm. Thus, ostensibly uninformative advertising directs market share to the lowest-cost supplier and promotes productive efficiency.

We now characterize the expected profit for firms in the advertising equilibrium. Using (2) and integrating by parts, we find that expected profit may be represented as:

$$
E_\theta\left[\Pi(\theta, \theta; A)\right] = r(p(\overline{\theta}), \overline{\theta})\frac{U}{N} + E_\theta\left[D(p(\theta))\frac{F}{f}(\theta)\left[\frac{U}{N} + [1 - F(\theta)]^{N-1}I\right]\right]. \tag{3}
$$

The first term on the RHS is the "profit at the top." The fully sorting scheme allocates the lowest market share for the type at the top: the probability of winning the informed consumers is zero for the highest type, $\overline{\theta}$. The second term represents the expected information rents. In regard to the magnitude of the second term, the fully sorting scheme has both a strength and a weakness. To see this, consider how the market share allocation affects the magnitude of the term. The strength of the fully sorting scheme is based on downward-sloping demand. Lower-cost firms set lower prices and thus generate greater demand from visiting consumers; hence, by directing more market share to lower-cost firms, the fully sorting scheme acts to expand the size of the market and increase expected information rents. The weakness of the fully sorting scheme is associated with the term $\frac{F}{f}(\theta)$. When greater market share is directed to type θ, this type earns greater profit and is thus

less tempted to mimic lower types. Lower types can then also earn greater profit without inducing a violation of incentive compatibility. Intuitively, the ratio $\frac{F}{f}(\theta)$ then describes the contribution to expected profit that is made when type θ receives greater market share, since this measures the proportion of types below θ conditional on the occurrence of type θ. Suppose that F is log-concave ($\frac{F}{f}(\theta)$ is nondecreasing in θ).[15] Then an increase in market share to type θ contributes more to expected profit when type θ is higher. The fully sorting scheme minimizes the market share that is allocated to higher types and thus works against the direction to which log-concavity of F appeals.

The advertising equilibrium can be understood as a purification of Bagwell and Ramey (1994a). In their paper, advertising directs market share to the firm that offers the best deals (in terms of price and variety) but equilibrium advertising takes the form of a mixed strategy. To see how our model constructs a purified version, consider a complete-information game, where production costs are fixed at a constant $c > 0$. Then, as we establish in the supplementary materials for this paper, there exists a unique symmetric mixed-strategy equilibrium in this game as in Bagwell and Ramey (1994a) and Varian (1980).[16] Consider next an incomplete-information game, where production costs rise in types θ. As we show in the supplementary materials, if a firm of type θ uses the advertising strategy $A(\theta)$ in the unique advertising equilibrium of the incomplete-information game, then the probability distribution induced by A is approximately the same as the distribution of advertising in the mixed-strategy equilibrium of the complete-information game, when the production costs for types θ (say, $c(\theta)$) approximate the constant c. This purification result offers a useful link between the complete- and incomplete-information analyses; however, it does not establish whether the main predictions of Bagwell and Ramey carry over when, as seems plausible, production costs vary significantly with types. As we show below, when some additional structure is placed on the demand and distribution functions, the main predictions of the complete-information model can be captured in the general incomplete-information setting.

2.3 Random Equilibrium

In this subsection, we analyze the random equilibrium, wherein all consumers use the random search rule and thus divide up evenly across firms. Each firm then receives an equal share, $\frac{1}{N}$, of the unit mass of consumers. Given the random search rule, firms necessarily choose zero advertising, since even informed consumers are unresponsive to advertising; furthermore, when firms pool and do not advertise, the random search rule is a best response for each consumer.[17] The random equilibrium

---

[15] This assumption is common in the contract literature and is satisfied by many distribution functions.

[16] In the complete-information game considered here, all firms set the same price and informed consumers are indifferent when using the advertising search rule. By contrast, Bagwell and Ramey (1994a) allow firms to make cost-reducing investments, and this ensures that higher-advertising firms offer strictly lower prices. In the analysis of advertising equilibria considered here, the advertising search rule is strictly optimal for informed consumers provided that incomplete information is present so that production costs vary (at least a little) with types.

[17] If informed consumers observe a deviation whereby some firm selects positive advertising, then random search remains optimal in the event that informed consumers believe that the deviating firm has an average type. Since such a deviation may be more attractive to a lower-cost type, the random equilibrium may fail to be a "refined" equilibrium in the static model. See Bagwell and Ramey (1994b) for an analysis of the refined equilibrium in a related model of advertising in which one firm has two possible cost types. As noted in the Introduction, the random equilibrium

thus exists and takes the form of a pooling equilibrium.

In the random equilibrium, the interim-stage profit for the firm of type θ is given by $r(p(\theta), \theta)\frac{1}{N}$. The random equilibrium sacrifices productive efficiency; however, all advertising expenses are avoided. Using $\frac{dr(p(\theta),\theta)}{d\theta} = -D(p(\theta))$, it is straightforward to confirm that the expected profit for a firm in the random equilibrium is

$$E_\theta \left[r(p(\theta), \theta)\frac{1}{N} \right] = r(p(\overline{\theta}), \overline{\theta})\frac{1}{N} + E_\theta \left[D(p(\theta))\frac{F}{f}(\theta)\frac{1}{N} \right]. \tag{4}$$

The RHS contains the profit at the top and the expected information rents, respectively.

2.4 Comparison of Advertising and Random Equilibria

We now compare the advertising and random equilibria. As illustrated in (3) and (4), in both types of equilibria, expected profit consists of two terms: the profit at the top and the expected information rents. To increase the profit at the top, the random equilibrium (pooling) is strictly preferred to the advertising equilibrium (full sorting). Intuitively, the highest-cost firm is never "out-advertised" in the random equilibrium and thus sells to its share of all consumers, $\frac{1}{N}$; by contrast, in the advertising equilibrium, the highest-cost firm is always out-advertised and thus sells only to its share of uninformed consumers, $\frac{U}{N}$. To increase expected information rents, however, it is not immediately clear whether the random or advertising equilibrium is preferred. On the one hand, if $\frac{F}{f}(\theta)$ is nondecreasing, then the random equilibrium is attractive, since this equilibrium allocates more market share to higher-cost types. On the other hand, downward-sloping demand creates a force that favors the advertising equilibrium, which allocates more market share to lower-cost types, since these types price lower and thus generate larger demand $D(p(\theta))$.

For the special case in which the support of possible cost types is small, we can unambiguously rank expected profits under the advertising and random equilibria. In particular, as $\overline{\theta} - \underline{\theta}$ approaches zero, expected information rents also approach zero in both the advertising and random equilibria. Profit at the top remains strictly higher under the random equilibrium, however, since the highest-cost firm gets strictly more market share in the random than the advertising equilibrium. Thus, for $\overline{\theta} - \underline{\theta}$ sufficiently small, expected profit is strictly higher under the random equilibrium than under the advertising equilibrium. Given the purification result described above and established in our supplementary materials, this finding can be understood as a direct extension of Bagwell and Ramey's (1994a) analogous finding for the associated complete-information game.

Consider next the general case in which the support of possible costs may be large. To go further in ranking expected profits, we must formally analyze the expected information rents.[18] Let A denote the advertising schedule used in the advertising equilibrium, in which the market

---

can also be associated with a setting in which advertising is prohibited (in which case deviant positive advertising selections are not possible). Our analysis here of random equilibria is also useful when we later consider the repeated game and the possibility of a self-enforcing agreement among firms in which a deviation from zero advertising would cause a future advertising war.

[18]Our analysis here builds on arguments made by Athey, Bagwell and Sanchirico (2004) in their analysis of price competition and collusion.

share allocation, $M(\theta; A) = \frac{U}{N} + [1 - F(\theta)]^{N-1}I$, is strictly decreasing. Similarly, let $A^p \equiv 0$ denote the advertising schedule used in the random (pooling) equilibrium, in which the market share allocation, $M(\theta; A^p) \equiv \frac{1}{N}$, is constant. We now define the distribution function

$$G(\theta; A) \equiv \frac{\int_{\underline{\theta}}^{\theta} M(x; A)f(x)dx}{\int_{\underline{\theta}}^{\overline{\theta}} M(x; A)f(x)dx}.$$

The distribution $G(\theta; A^p)$ is similarly defined. The denominator represents the (ex ante) expected market share, which equals $\frac{1}{N}$. Since $M(\theta; A)$ is strictly decreasing, $M(\theta; A^p) = \frac{1}{N}$ crosses $M(\theta; A)$ from below. This implies in turn that $G(\theta; A^p)$ first-order stochastically dominates $G(\theta; A)$: $G(\theta; A^p) \le G(\theta; A)$. Thus, if $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing, then

$$\int_{\underline{\theta}}^{\overline{\theta}} D(p(\theta))\frac{F}{f}(\theta)dG(\theta; A^p) \ge \int_{\underline{\theta}}^{\overline{\theta}} D(p(\theta))\frac{F}{f}(\theta)dG(\theta; A).$$

The inequality can be rewritten as

$$E_{\theta}\left[D(p(\theta))\frac{F}{f}(\theta)M(\theta; A^p)\right] \ge E_{\theta}\left[D(p(\theta))\frac{F}{f}(\theta)M(\theta; A)\right]. \tag{5}$$

Referring to (3)-(5), we conclude that, if $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing, then expected information rents are weakly higher in the random equilibrium than in the advertising equilibrium.

Summarizing, in comparison to the advertising equilibrium, the random equilibrium has strictly higher profit at the top and, if $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing, weakly higher expected information rents. As suggested above, $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing if the log-concavity of F is significant in comparison to the extent to which demand slopes down. Further insight is possible by considering the limiting case in which $D(p(\theta))$ is perfectly inelastic, so that $D(p(\theta))$ is constant for all prices up to a reservation value. In this case, if $F(\theta)$ is log-concave, then $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing.[19]

We may now state the following conclusion:

Proposition 2. *If F is log-concave and demand is sufficiently inelastic, or if the support of possible cost types is sufficiently small, then firms make a strictly higher expected profit in the random equilibrium than in the advertising equilibrium.*

Proposition 2 indicates that important circumstances exist under which firms gain when the use of advertising is restricted. As our discussion of the random equilibrium confirms, advertising would not be used if informed consumers were to ignore it. If informed consumers were responsive to advertising, however, then firms might nevertheless achieve a restriction on the use of advertising if advertising were legally prohibited. For a fixed industry structure, Proposition 2 thus suggests that retail firms would benefit from a prohibition on non-price retail advertising. A further possibility is that firms are able to eliminate the use of advertising through a self-enforcing collusive agreement

---

[19]In fact, if demand is perfectly inelastic and F is log-concave, the random equilibrium generates strictly higher expected information rents than the advertising equilibrium. This follows since $\frac{F}{f}(\theta)$ is strictly increasing at $\underline{\theta}$.

and that firms prefer such a restriction to *any* other self-enforcing advertising scheme. We delay further consideration of this possibility until the next section.

Proposition 2 establishes that firms gain by restricting the use of advertising if F is log-concave and demand is sufficiently inelastic or if the support of possible cost types is sufficiently small. It is important to note, though, that this conclusion may hold even when the assumptions are weakened. Consider the CES demand function, $D(p) = p^{-\epsilon}$, and suppose that demand is elastic (i.e., $\epsilon > 1$). Assume further that F is log-concave in the specific sense that types are distributed uniformly over $[\underline{\theta}, \overline{\theta}]$ where $\underline{\theta} > 0$. For this example, calculations reveal that $\frac{d}{d\theta}[D(p(\theta))\frac{F}{f}(\theta)] > 0$ if $\overline{\theta}/[\overline{\theta} - \underline{\theta}] > \epsilon$. Firms thus earn a strictly higher expected profit by pooling at zero advertising than by following the advertising equilibrium, provided that the elasticity of demand, ϵ, does not exceed a critical level where this level is higher when the support of possible cost types is smaller.

2.5   Free-Entry Equilibrium

We now relax the assumption that the number of firms is fixed. To this end, following Bagwell and Ramey (1994a), we include now an initial stage for the game in which firms simultaneously decide whether to enter, where entry entails a positive setup (or opportunity) cost. After a firm chooses to enter, it privately learns its cost type. The number of entering firms is publicly observed, and the game then proceeds as above.

It is clear from (3) and (4) that expected profit is strictly decreasing in the number of firms, N, whether firms anticipate the advertising or random equilibrium. Thus, in each case, an equilibrium number of firms is implied such that the profit from entry (inclusive of the fixed cost) would be negative were one more firm to enter. Let N^s denote the equilibrium number of entering firms when the advertising (full sorting) equilibrium is anticipated, and let N^p denote the equilibrium number of entering firms when the random (pooling) equilibrium is expected. It is also clear from Proposition 2 that, if F is log-concave and demand is sufficiently inelastic, or if $\overline{\theta} - \underline{\theta}$ is sufficiently small, then $N^p \geq N^s$. Under these conditions, at least as many firms enter when the random equilibrium is expected as when the advertising equilibrium is anticipated.

The model also leads to welfare comparisons. Assume that $\min(N^s, N^p) \geq 1$.[20] When the number of firms is endogenized, if we ignore integer constraints, then firms earn zero expected profit whether the random or advertising equilibrium is anticipated. Uninformed consumers are also indifferent. Intuitively, under either equilibrium, an uninformed consumer picks a firm at random and thus faces an expected price of $E_\theta p(\theta)$. Finally, consider the informed consumers. When the random equilibrium occurs, an informed consumer also faces an expected price of $E_\theta p(\theta)$; however, when the advertising equilibrium occurs, an informed consumer is guided by advertising activity to the lowest market price and thus faces the expected *minimum* price in the market. Provided that $N^s \geq 2$, an informed consumer thus strictly prefers the advertising equilibrium. When the number of firms is endogenous, it follows that expected welfare is higher when the advertising equilibrium is anticipated than when the random equilibrium is expected. This conclusion does not require any

---

[20] When $N = 1$, a single firm enters the market and chooses $A = 0$, and all consumers visit it.

assumption as to the elasticity of demand or the log-concavity of the distribution function.

We may now summarize with the following proposition:

Proposition 3. *Assume that* $\min(N^s, N^p) \geq 1$. *(i) If* F *is log-concave and demand is sufficiently inelastic, or if the support of possible cost types is sufficiently small, then* $N^p \geq N^s$ *(concentration is at least as high in the advertising equilibrium as in the random equilibrium). (ii) Social surplus is as high in the advertising equilibrium as in the random equilibrium; further, if* $N^s \geq 2$, *then social surplus is strictly higher in the advertising equilibrium than in the random equilibrium.*

Allowing that the support of possible costs may be large, we thus establish a general sense in which Bagwell and Ramey's main findings extend to the private-information setting. When legal or other considerations lead to the absence of advertising, if the distribution of types is log-concave and demand is sufficiently inelastic, then the market is less concentrated than it would be were advertising competition to occur. Furthermore, the average transaction price is lower, and social welfare is thus higher, when entry is endogenized and firms compete in advertising. Note, however, that some findings such as Proposition 2 and Proposition 3 (i) are not straightforward, given downward-sloping demand. For a given number of firms, pooling at zero advertising acts to increase the profit at the top but sorting through advertising acts to increase expected information rents when demand is substantially larger for lower prices. This conflict suggests that market concentration could be lower in the advertising equilibrium than in the random equilibrium, when demand is sufficiently elastic. Thus, the established positive association between advertising and market concentration employs additional assumptions on the distribution of types and the elasticity of demand in the general private-information setting.

It is interesting to compare these findings with empirical patterns emphasized in the earlier literature on advertising. Benham (1972) provides evidence for retail markets that prices are lower and market concentration is higher, when non-price retail advertising is allowed. Our findings offer theoretical support for these associations. In another set of studies, Bain (1956), Comanor and Wilson (1974) and others find a positive relationship between manufacturer advertising and profitability. These authors suggest that the relationship may reflect the role of advertising in deterring entry. Consistent with interpretations offered by Demsetz (1973) and Nelson (1974), our work suggests that advertising and profitability may be positively related, since they are both implications of superior efficiency. In particular, in the advertising equilibrium, lower-cost firms advertise more, have larger sales and earn greater profit.

2.6 Comparison with Pricing Equilibrium

In this subsection, we compare the advertising equilibrium with the analogous pricing equilibrium that emerges in a benchmark model in which $N \geq 2$ ex ante identical firms compete in prices. In particular, we follow Varian (1980) and suppose that informed consumers observe prices and buy from the lowest-priced firm while uninformed consumers pick a firm at random. Following Spulber (1995) and Bagwell and Wolinsky (2002), we modify Varian's model and allow that firms

are privately informed as to their costs. We characterize the *pricing equilibrium* of this benchmark game and compare the associated expected profit with that achieved in the advertising equilibrium of our static game.

In the benchmark game, if a pricing strategy is denote by ρ, then the interim-stage profit in direct form is given by

$$\Pi^B(\widehat{\theta}, \theta; \rho) = [\rho(\widehat{\theta}) - \theta]D(\rho(\widehat{\theta}))M^B(\widehat{\theta}; \rho),$$

where we use the superscript B to denote the benchmark (Bertrand) game. When a firm selects the price $\rho(\widehat{\theta})$ and other firms use the pricing strategy ρ, then the firm's expected market share is denoted as $M^B(\widehat{\theta}; \rho)$. The profit-if-win is defined by $[\rho(\widehat{\theta}) - \theta]D(\rho(\widehat{\theta})) \equiv r(\rho(\widehat{\theta}), \theta)$. As in Spulber (1995), a unique and symmetric equilibrium can be established. A new feature in our benchmark model is that uninformed consumers exist. The pricing equilibrium ρ satisfies:

$$\rho'(\theta) = -\frac{r(\rho(\theta), \theta)[\partial M^B(\theta; \rho)/\partial \theta]}{r_\rho(\rho(\theta), \theta)M^B(\theta; \rho)} \text{ and } \rho(\overline{\theta}) = p(\overline{\theta}), \tag{6}$$

where $M^B(\theta; \rho) = \frac{U}{N} + [1 - F(\theta)]^{N-1}I$. Straightforward arguments ensure that the equilibrium price is lower than the monopoly price except the price at the top, so that $r_\rho > 0$. As (6) confirms, the equilibrium pricing schedule is strictly increasing; thus, firms are fully sorted by their types in the pricing equilibrium. Notice that the highest-cost firm selects its monopoly price, $p(\overline{\theta})$, and sells only to uninformed consumers.

In the pricing equilibrium, interim-stage profit can be written as

$$\Pi^B(\theta, \theta; \rho) = \Pi^B(\overline{\theta}, \overline{\theta}; \rho) + \int_\theta^{\overline{\theta}} D(\rho(x))\left[\frac{U}{N} + [1 - F(x)]^{N-1}I\right]dx, \tag{7}$$

where the profit at the top is $\Pi^B(\overline{\theta}, \overline{\theta}; \rho) = r(p(\overline{\theta}), \overline{\theta})\frac{U}{N}$. Integrating by parts, we find that expected profit is given as:

$$E_\theta\left[\Pi^B(\theta, \theta; \rho)\right] = r(p(\overline{\theta}), \overline{\theta})\frac{U}{N} + E_\theta\left[D(\rho(\theta))\frac{F}{f}(\theta)\left[\frac{U}{N} + [1 - F(\theta)]^{N-1}I\right]\right]. \tag{8}$$

Comparing (8) with (3), we see that the profit at the top is the same in the advertising equilibrium as in the pricing equilibrium. In each case, the highest-cost firm monopolizes only uninformed consumers. The expected information rents are higher in the pricing equilibrium, however, since demand is greater when prices are set below monopoly levels. We thus have the following conclusion: for any fixed number of firms, a firm's expected profit is strictly higher in the pricing equilibrium than in the advertising equilibrium.[21] Evidently, when firms possess private information about their costs, competition in (non-price) advertising is more aggressive than (Bertrand) competition in prices. Intuitively, price competition induces greater in-store demand from consumers and thus

---

[21]In a different context, Bagwell and Ramey (1988) present a somewhat related finding. Working with a two-type signaling model, they show that a low-cost incumbent earns greater profit when it separates using price as a signal than when it separates using wasteful advertising (money-burning) as a signal.

elevates the size of expected information rents for firms. When the number of firms is fixed, both consumers and firms agree that the pricing equilibrium is preferred to the advertising equilibrium. When the number of firms is endogenized by the free-entry condition, more firms enter in the former equilibrium than in the latter equilibrium. Once market structure is endogenized, firms are indifferent between pricing and advertising competition, but consumers strictly prefer the former to the latter (provided that at least two firms enter in the pricing equilibrium).

We may thus summarize the findings of this subsection as follows:

Proposition 4. *There exists a unique and symmetric pricing equilibrium, and in this equilibrium the pricing function $\rho(\theta)$ satisfies $\rho(\theta) > \theta$ and is strictly increasing and differentiable. Expected profit and consumer surplus are both strictly higher in the pricing equilibrium than in the advertising equilibrium. Further, when the number of firms is endogenized, at least as many firms enter in the pricing equilibrium as in the advertising equilibrium; and, if at least two firms enter in the pricing equilibrium, then social surplus is strictly higher in the pricing equilibrium than in the advertising equilibrium.*

With these findings at hand, we may now offer a further interpretation of Benham's findings. Let us associate the advertising equilibrium with a setting in which only non-price advertising is allowed, the pricing equilibrium with a setting in which price advertising is allowed, and the random equilibrium with a setting in which advertising is banned. Provided that the market always has at least two firms, our results in this section indicate that the average transaction price is lowest when price advertising is allowed, somewhat higher when only non-price advertising is allowed, and higher yet when all advertising is banned. Likewise, when the number of firms is endogenous, social welfare is highest when price advertising is allowed, somewhat lower when only non-price advertising is allowed, and lower yet when all advertising is banned. Finally, when demand is sufficiently inelastic and the distribution of types is log-concave, the market is less concentrated when advertising is banned than when non-price or price advertising is allowed.[22] These findings are broadly consistent with Benham's findings.

2.7 Comparative Statics

We now return to the static model of advertising and conduct comparative-statics analysis. In particular, we consider how the advertising equilibrium responds to changes in the parameters, I and N, and to shifts of the distribution function of types.

To analyze comparative statics associated with distribution functions, we consider distribution functions F and G that have the same support $[\underline{\theta}, \overline{\theta}]$. As above, the distribution functions are twice-continuously differentiable and have positive densities f and g. We then compare two advertising

---

[22]For a fixed number of firms, if demand is perfectly inelastic, the expected information rents in the pricing equilibrium are the same as in the advertising equilibrium. Thus, when demand is sufficiently inelastic, market concentration is approximately the same in these two equilibria. Further, if $\overline{\theta} - \underline{\theta}$ is sufficiently small, then the market is less concentrated when advertising is banned than when non-price or price advertising is allowed. This is because the random equilibrium generates the largest market share for a firm with cost type $\overline{\theta}$.

equilibrium strategies, $A_F(\theta)$ and $A_G(\theta)$, that correspond to the distribution functions, F and G, respectively. We compare the distributions F and G by using the monotone likelihood ratio (MLR) order. The distribution function F dominates G in terms of the MLR order if $\frac{f(\theta)}{g(\theta)}$ is strictly increasing for all $\theta \in [\underline{\theta}, \overline{\theta}]$. Using the MLR order, we can show how firms choose their advertising when lower-cost (higher-advertising) types are more likely under G than under F.

Our comparative-statics results are contained in the following proposition:

Proposition 5. *(i) Equilibrium advertising $A(\theta)$ is strictly increasing in I for all $\theta < \overline{\theta}$ with $A(\overline{\theta}) = 0$. (ii) If N rises, then there exists $\widehat{\theta} \in (\underline{\theta}, F^{-1}(1 - e^{-\frac{1}{N-1}}))$ such that equilibrium advertising strictly increases for $\theta \in [\underline{\theta}, \widehat{\theta})$, strictly decreases for $\theta \in (\widehat{\theta}, \overline{\theta})$, and is unchanged for $\theta \in \{\widehat{\theta}, \overline{\theta}\}$. (iii) For all $\theta > \underline{\theta}$ and $\varepsilon > 0$, there exists N' such that, for all $N > N'$, $A(\theta) < \varepsilon$. (iv) If distribution function F dominates G in terms of the MLR order, then there exists $\widetilde{\theta} \in (\underline{\theta}, \overline{\theta})$ such that $A_F(\theta) < A_G(\theta)$ for $\theta \in [\underline{\theta}, \widetilde{\theta})$, $A_F(\theta) > A_G(\theta)$ for $\theta \in (\widetilde{\theta}, \overline{\theta})$, and $A_F(\theta) = A_G(\theta)$ for $\theta \in \{\widetilde{\theta}, \overline{\theta}\}$.*

The proofs of parts (ii) and (iv) are in the Appendix.

Using the derivation of $A(\theta)$ in the proof of Proposition 1, we can immediately confirm that part (i) holds.[23] Intuitively, firms compete more intensely by raising advertising when the gain from capturing informed consumers rises. It is less clear, however, whether advertising increases when N rises. On the one hand, an increase in the number of firms may lead to greater competition for the informed consumers and thus an increase in advertising. On the other hand, an increase in the number of firms may also cause firms to become discouraged about the prospect of winning the informed consumers and thus result in a decrease in advertising. In part (ii), we confirm that these competing considerations weigh differently across firms with different cost types: when the number of firms increases, lower-cost firms compete more aggressively and raise advertising, while higher-cost firms perceive a reduced chance of winning the informed consumers and lower advertising. An interesting implication is that the support of equilibrium advertising levels (i.e., $[A(\overline{\theta}) = 0, A(\underline{\theta})]$) is larger in markets with more firms. Observe, however, that as the number of firms goes to infinity, the cutoff type $\widehat{\theta}$ converges to $\underline{\theta}$; thus, for markets with a sufficiently large number of firms, further entry is almost sure to lower the advertising of any given firm. In fact, we can easily confirm that part (iii) holds and thus that, for any type other than the lowest type, the equilibrium level of advertising must be near zero when the number of firms is sufficiently large.[24]

Finally, as part (iv) establishes, competing considerations arise as well when the distribution of costs changes so that lower-cost realizations become more likely in the sense of the MLR order. Following such a shift, lower-cost firms compete more aggressively for informed consumers and thus increase their advertising; however, higher-cost firms become discouraged about their chances of winning the informed consumers and thus lower their advertising levels. Our work here builds on Hopkins and Kornienko (2007), who report a similar finding for a family of all-pay auctions.

---

[23] Formally, this follows since $\partial M(x; A)/\partial x$ is strictly decreasing in I for all $x < \overline{\theta}$.

[24] As shown in the proof of Proposition 1, $A(\overline{\theta}) = 0$ and $A'(\theta) = r(p(\theta), \theta)\partial M(\theta; A)/\partial \theta$. Part (iii) thus follows, since, for all $\theta > \underline{\theta}$, $\partial M(\theta; A)/\partial \theta$ goes to zero as N goes to infinity.

3 The Repeated Game

If the number of firms is fixed at N and firms can collude, what would they do? Our findings in the static model suggest that firms may prefer a situation in which they select zero advertising (or the minimal advertising that ensures consumers know the firms exist). The random equilibrium can be achieved in a static setting if informed consumers ignore advertising; alternatively, a legal ban on advertising can enable firms to eliminate advertising competition. Putting these possibilities to the side, we assume henceforth that informed consumers use the advertising search rule and that advertising is legal. Even under these assumptions, if firms interact repeatedly through time, they may limit the use of advertising as part of a self-enforcing agreement. We are thus led to consider a repeated game in which firms are privately informed with respect to their cost levels. In this section, we define the repeated game and present some programs that are useful in the next section where we characterize optimal collusion.

3.1 The Model

We now define the repeated game. In each of an infinite number of periods, firms play the static game defined in Section 2. We assume henceforth that, in each period, informed consumers use the advertising search rule. Uninformed consumers again use the random search rule. As shown in Section 2, these search rules are optimal in a given period if firms use symmetric strategies and lower-cost types always advertise at (weakly) higher levels. As discussed in more detail below, for the equilibrium concept that we employ, these requirements for firms' strategies are satisfied. Hence, in our formal definitions of the repeated game and the equilibrium concept, we may simplify and focus exclusively on the behavior of firms.

Upon entering a period, firms share a public history, in that each firm observes the realized advertising expenditures of all firms in all previous periods. A firm also privately observes its current cost type. As well, each firm privately observes the history of the cost types that it had, the prices that it selected and the advertising schedules that it used in previous periods. Thus, we consider a setting in which a firm does not observe any rival firm's current or past cost types and also does not observe any rival firm's current or past advertising schedules. In addition, a firm does not observe the realized price choice of any rival in any past period.[25]

The vectors of cost types, advertising schedules and realized advertisements at date t are denoted $\boldsymbol{\theta}_t \equiv (\theta_{it}, \theta_{-it})$, $\mathbf{A}_t \equiv (A_{it}, A_{-it})$ and $\mathbf{a}_t \equiv (a_{it}, a_{-it})$. Under the assumed consumer search rules, let $m_i(\mathbf{a}_t)$ denote the market share received by firm i when the advertising vector $\mathbf{a}_t$ is used. Then, an infinite sequence $\{\boldsymbol{\theta}_t, \mathbf{A}_t\}_{t=1}^{\infty}$ generates a path-wise payoff for firm i:

$$u_i(\{\boldsymbol{\theta}_t, \mathbf{A}_t\}_{t=1}^{\infty}) = \sum_{t=1}^{\infty} \delta^{t-1} \left[r\left(p\left(\theta_{it}\right), \theta_{it}\right) m_i(\mathbf{a}_t) - a_{it} \right].$$

---

[25] In the supplementary materials for this paper, we discuss the implications of relaxing this assumption. Our main results are robust to this relaxation, if demand is sufficiently inelastic.

Notice that we embed the monopoly price selection into the net revenue function, r. This simplifies the analysis and is without loss of generality given our assumption that past prices are not public among firms. As in the static model, we assume that cost shocks are iid across firms. For the repeated game, we introduce as well the assumption that cost shocks are iid over time. With this assumption, the repeated game takes a recursive structure.

As our solution concept, we employ Perfect Public Equilibrium (Fudenberg, Levine and Maskin, 1994). We thus focus on public strategies. A firm uses a public strategy when a firm's current advertising level depends on its current cost level and the public history of realized advertising levels. At the close of date τ, the public history of realized advertisements is $h_\tau = \{\mathbf{a}_t\}_{t=1}^{\tau}$. Let H_τ be the set of potential public histories at date τ. A public strategy for firm i in period τ, $s_{i\tau}$, is a mapping from $H_{\tau-1}$ to the set of stage-game strategies $\{A \mid A : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}_+\}$. A public strategy for firm i, s_i, is then a sequence $\{s_{it}\}_{t=1}^{\infty}$, and a profile of public strategies is $\mathbf{s} = \{s_1, ..., s_N\}$. We restrict attention to *Symmetric Perfect Public Equilibrium (SPPE)*, whereby $s = s_1 = ... = s_N$. Thus, in an SPPE, firms adopt symmetric advertising schedules after every history: $s_{i\tau}(h_{\tau-1}) = s_{j\tau}(h_{\tau-1})$ for all i, j, τ and $h_{\tau-1}$.

3.2   Dynamic Programming Approach

Building on work by Abreu, Pearce and Stacchetti (1986, 1990) [APS], we apply a dynamic programming approach to our recursive setting. Let $V \subset \mathbb{R}$ be the set of SPPE values. Note that, at this point, we have not established $\sup V \in V$ or $\inf V \in V$. Following APS, any symmetric public strategy profile $\mathbf{s} = \{s, ..., s\}$ can be factored into two components: a first-period advertising schedule A and a continuation-value function $v : \mathbb{R}_+^N \to \mathbb{R}$. The continuation-value function describes the repeated-game expected payoff enjoyed by all firms as evaluated at the beginning of period two, before period-two cost types are realized. This payoff is allowed to depend on the first-period advertising realization $\mathbf{a} \equiv (a_1, ..., a_N) \in \mathbb{R}_+^N$.

Under this approach, for any given symmetric public strategy profile $\mathbf{s}$, we may ignore subscript i (as in the static model) and denote the interim-stage first-period profit for firm i of type θ as $\Pi(A(\theta), \theta; A) \equiv R(A(\theta), \theta; A) - A(\theta)$. At the interim-stage in the first period, firm i's expected continuation value may be denoted as $\overline{v}(A(\theta); A) \equiv E_{\theta_{-i}}[v(A(\theta), A_{-i}(\theta_{-i}))]$, where $A_{-i}(\theta_{-i})$ denotes the $(N-1)$-tuple of advertising selections by other firms when these firms all use the schedule A. Letting $\delta \in (0, 1)$ denote the common discount factor for firms, we may now use $\Pi(A(\theta), \theta; A) + \delta \overline{v}(A(\theta); A)$ to represent a firm's interim-stage payoff from a symmetric public strategy profile $\mathbf{s}$. A firm's expected payoff from $\mathbf{s}$ is then given as $E_\theta [\Pi(A(\theta), \theta; A) + \delta \overline{v}(A(\theta); A)]$.

The set of optimal SPPE can be characterized by solving a "factored program". In particular, we may choose an advertising schedule and a continuation-value function to maximize the expected payoff to a firm subject to feasibility and incentive constraints.

Factored Program: The program chooses an advertising schedule A and a continuation-value function v to maximize

$$E_\theta [\Pi(A(\theta), \theta; A) + \delta \overline{v}(A(\theta); A)]$$

subject to: (i) for all $\mathbf{a}$, $v(\mathbf{a}) \in V$, and (ii) for any deviation $\widehat{A}$,

$$E_\theta \left[\Pi(A(\theta), \theta; A) + \delta \overline{v}(A(\theta); A) \right] \geq E_\theta [\Pi(\widehat{A}(\theta), \theta; A) + \delta \overline{v}(\widehat{A}(\theta); A)].$$

A key implication of the dynamic programming approach is that the set of optimal SPPE can be characterized by solving the Factored Program. Specifically, let $\mathbf{s}^* = \{s^*, ..., s^*\}$ be a symmetric public strategy profile with the corresponding factorization (A^*, v^*). Then, s^* is an optimal SPPE if and only if (A^*, v^*) solves the Factored Program.

We next follow Athey and Bagwell (2001) and Athey, Bagwell and Sanchirico (2004), who show that existing tools from (static) mechanism design theory can be used to find the optimal factorization. To this end, we rewrite the Factored Program as an Interim Program. The latter program utilizes interim-stage profit and parses the incentive constraint into two kinds: (i) the "on-schedule" constraint that each firm truthfully announces its cost and (ii) the "off-schedule" constraint that each firm cannot gain by choosing an advertising level that is not assigned to any cost type.

Interim Program: The program chooses A and v to maximize

$$E_\theta \left[\Pi(A(\theta), \theta; A) + \delta \overline{v}(A(\theta); A) \right]$$

subject to:
(i) On-schedule incentive compatibility: $\forall \widehat{\theta}$,

$$\forall \theta_{-i}, v(A(\widehat{\theta}), A_{-i}(\theta_{-i})) \in V$$
$$\forall \theta, \ \Pi(A(\theta), \theta; A) + \delta \overline{v}(A(\theta); A) \geq \Pi(A(\widehat{\theta}), \theta; A) + \delta \overline{v}(A(\widehat{\theta}); A)$$

(ii) Off-schedule incentive compatibility: $\forall \widehat{a} \notin A([\underline{\theta}, \overline{\theta}])$,

$$\forall \theta_{-i}, \ v(\widehat{a}, A_{-i}(\theta_{-i})) \in V$$
$$\forall \theta, \ \Pi(A(\theta), \theta; A) + \delta \overline{v}(A(\theta); A) \geq \Pi(\widehat{a}, \theta; A) + \delta \overline{v}(\widehat{a}; A)$$

Following Athey, Bagwell and Sanchirico (2004), we next relax the Interim Program in two ways. First, we ignore the off-schedule constraints by assuming that δ is sufficiently high so that no off-schedule deviation is profitable. Second, we relax the on-schedule constraints by replacing $v(A(\widehat{\theta}), A_{-i}(\theta_{-i})) \in V$ with $\overline{v}(A(\widehat{\theta}); A) \leq \sup V$. The relaxed constraint thus requires only that the expected continuation value does not exceed the supremum of SPPE. When the constraints are relaxed in this way, we have the Relaxed Program.

To facilitate connection with tools from mechanism design theory, we next re-write the Relaxed Program using direct-form notation. Formally, as above, we let $\Pi(\widehat{\theta}, \theta; A) \equiv \Pi(A(\widehat{\theta}), \theta; A)$ and $R(\widehat{\theta}, \theta; A) \equiv R(A(\widehat{\theta}), \theta; A)$. We also define $W(\widehat{\theta}) \equiv \delta[\sup V - \overline{v}(A(\widehat{\theta}); A)]$. For instance, $W(\widehat{\theta}) > 0$ means that the expected continuation value falls below the value $\sup V$ subsequent to a firm's

21

announcement $\widehat{\theta}$. A continuation-value reduction represents a "war" that involves an increase of advertising expenses in the future. We may now state the Relaxed Program in terms of the choice of the current-period advertising schedule A and the "punishment" function W that maximizes expected payoff subject to on-schedule constraints:

Relaxed Program: The program chooses A and W to maximize

$$E_\theta\left[R(\theta, \theta; A) - A(\theta) - W(\theta)\right]$$

subject to:

$$\forall\theta,\ W(\theta) \geq 0$$
$$\text{(On-IC)}\ \forall\theta,\widehat{\theta},\ R(\theta, \theta; A) - A(\theta) - W(\theta) \geq R(\widehat{\theta}, \theta; A) - A(\widehat{\theta}) - W(\widehat{\theta}).$$

To see that the Relaxed Program is indeed a relaxation of the Interim Program, suppose that (A, v) satisfies the constraints of the Interim Program. Let us now translate (A, v) into (A, W) via $W(\widehat{\theta}) \equiv \delta[\sup V - \overline{v}(A(\widehat{\theta}); A)]$. Using this translation, it is now easy to confirm that (A, W) satisfies the constraints of the Relaxed Program and that the Interim and Relaxed Programs rank factorizations (A, v) in the same way. Therefore, if we find a solution (A, W) to the Relaxed Program, and if that solution can be expressed as a translation of some (A, v) that satisfies all of the constraints of the Interim Program, then this (A, v) is the factorization of an optimal SPPE.

Our next step is to identify an important situation in which the solution to the Relaxed Program can be translated back into an optimal SPPE factorization.

Proposition 6 (Stationarity): *Suppose that $(A^*, W^* \equiv 0)$ solves the Relaxed Program. Then there exists $\widehat{\delta} \in (0, 1)$ such that, for all $\delta \geq \widehat{\delta}$, there exists an optimal SPPE which is stationary, wherein firms use A^* after all equilibrium-path histories, and A^* solves the following program: maximize $E_\theta[R(\theta, \theta; A) - A(\theta)]$ subject to $\forall\theta,\widehat{\theta},\ R(\theta, \theta; A) - A(\theta) \geq R(\widehat{\theta}, \theta; A) - A(\widehat{\theta})$.*

To prove this proposition, we follow the steps used in the proof of Proposition 2 in Athey, Bagwell and Sanchirico (2004). In particular, we note two implications of the assumption that $(A^*, W^* \equiv 0)$ solves the Relaxed Program. First, following the discussion just above, $(A^*, v^* \equiv \sup V)$ is then a solution to the Interim Program, provided that this factorization satisfies the additional constraints of the Interim Program. We may therefore conclude that $(A^*, v^* \equiv \sup V)$ achieves a (weakly) higher payoff than can be achieved by any SPPE factorization. Thus, $E_\theta\left[\Pi(\theta, \theta; A^*) + \delta \sup V\right] \geq \sup V$. Second, if firms are sufficiently patient, then the repeated play of A^* in each period along the equilibrium path, with appropriate punishments off the equilibrium path, is in fact an SPPE. Given that $W^* \equiv 0$, A^* satisfies (IC-On) on a period-by-period basis. Likewise, A^* satisfies the on-schedule incentive constraint of the Interim Program on a period-by-period basis (i.e., when the continuation value does not vary with the on-schedule advertising level). The off-schedule incentive constraint of the Interim Program is also satisfied, provided that δ is sufficiently high. Repeated play of the (noncooperative) advertising equilibrium of the static game is always an SPPE of the

repeated game and may be used as the punishment that follows any off-schedule deviation.[26] Thus, when δ is sufficiently high, $E_\theta\left[\Pi(\theta, \theta; A^*)\right]/(1-\delta) \leq \sup V$. Using the two inequalities, we conclude that the repeated play of A^* is then an optimal SPPE: $\sup V = E_\theta[\Pi(\theta, \theta; A^*)]/(1-\delta)$.

Hence, if a solution of the Relaxed Program is $(A^*, W^* \equiv 0)$, and thus does not involve wars (i.e., is stationary), and if firms are sufficiently patient, then $\sup V$ is in fact in V. Further, an associated optimal SPPE can be easily characterized. Firms simply use the schedule A^* in each period, where A^* is the solution to the static program presented in Proposition 6. This result guides our subsequent analysis. Below, we use mechanism-design tools to characterize the (A, W) pairs that satisfy (On-IC) in the Relaxed Program. In the next section, we show that $(A^*, W^* \equiv 0)$ is always a solution to the Relaxed Program, and we also characterize A^*.

Consider now (On-IC) from the Relaxed Program. As the following lemma indicates, this constraint may be stated in a more useful way.

Lemma 1. (A, W) *satisfies on-schedule incentive compatibility (On-IC) if and only if* $\forall \theta$ *(i)* $A(\theta)$ *is nonincreasing and (ii)*

$$R(\theta, \theta; A) - A(\theta) - W(\theta) = R(\overline{\theta}, \overline{\theta}; A) - A(\overline{\theta}) - W(\overline{\theta}) + \int_\theta^{\overline{\theta}} D(p(x))M(x; A)dx. \qquad (9)$$

The proof of this result is standard in the mechanism-design literature and is therefore omitted.[27] The lemma indicates that the interim-stage expected payoff for a firm with period-one type θ is comprised of a payoff-at-the-top expression (i.e., $R(\overline{\theta}, \overline{\theta}; A) - A(\overline{\theta}) - W(\overline{\theta})$) and an integral that indicates the expected information rents for this type in the first period.

The repeated game allows for a wide range of behaviors, even within the category of stationary SPPE. For example, as noted, in each period of the repeated game, firms may use the advertising equilibrium of the static model. Further, under the conditions given in Proposition 2, firms strictly prefer pooling at zero advertising to using the advertising equilibrium of the static game. Hence, if those conditions hold and firms are sufficiently patient, then they can enforce a stationary SPPE in which they pool with zero advertising. Any pooling arrangement trivially satisfies on-schedule incentive compatibility, and patient firms will not deviate (off schedule) to a positive advertising level if such a deviation induces a future war that entails a reversion to the advertising equilibrium. Likewise, under appropriate conditions, stationary SPPE exist in which firms use advertising schedules that are nonincreasing step functions. More generally, stationary SPPE may entail advertising schedules with intervals of pooling as well as intervals of separation.

---

[26] We show below in Lemma 3 that A^* achieves strictly higher expected profit than does the advertising equilibrium of the static game.

[27] As in the proof of Proposition 1, it is straightforward to confirm that (On-IC) implies that $M(\theta; A)$ is nonincreasing. Given the consumer search rules, $M(\theta; A)$ is nonincreasing if and only if $A(\theta)$ is nonincreasing. A local optimality condition must also hold, and the application of an appropriate envelope theorem (Milgrom and Segal, 2002) thus yields (9). Together, the two conditions are sufficient for (On-IC), due to the single-crossing property of the model.

4   Optimal Collusion

In this section, we characterize optimal SPPE, assuming firms are sufficiently patient so that off-schedule constraints hold. First, we show that equilibrium-path wars are not necessary in an optimal SPPE. We thus show that an optimal SPPE exists that is stationary. Second, using Proposition 6, we report conditions under which an optimal SPPE involves pooling at zero advertising. Third, in a more general setting, we show that an optimal SPPE involves at least partial pooling. Fourth, we characterize the critical discount factor above which off-schedule constraints hold.

4.1   No Wars

We now show that equilibrium-path wars are not necessary in an optimal SPPE for patient firms. Suppose that a scheme (A, W) satisfies (On-IC) in the Relaxed Program. Then, we say that an alternative scheme $(\widetilde{A}, \widetilde{W})$ is *point-wise equivalent* to (A, W) if the scheme satisfies (On-IC) and preserves the market-share schedule and interim-stage profit:

$$\forall \theta, \ M(\theta; \widetilde{A}) = M(\theta; A) \text{ and } R(\theta, \theta; \widetilde{A}) - \widetilde{A}(\theta) - \widetilde{W}(\theta) = R(\theta, \theta; A) - A(\theta) - W(\theta).$$

Lemma 2. *Assume that (A, W) satisfies (On-IC) in the Relaxed Program. (i) There exists a no-wars scheme $(\widetilde{A}, \widetilde{W} \equiv 0)$ that is point-wise equivalent to (A, W). (ii) Any no-wars scheme $(\widetilde{A}, \widetilde{W} \equiv 0)$ that is point-wise equivalent to (A, W) satisfies $\widetilde{A}(\theta) \equiv A(\theta) + W(\theta)$.*

   Proof. The proof for (ii) is immediate by the definition of point-wise equivalence. Given $M(\theta; \widetilde{A}) = M(\theta; A)$, $R(\theta, \theta; \widetilde{A}) = R(\theta, \theta; A)$ follows; thus, $\widetilde{A}(\theta) = A(\theta) + W(\theta)$ must hold.

   To prove (i), we decompose the market-share allocation of (A, W) into three components: sorting intervals, pooling intervals and jump points. We then show that the intervals on which the no-war scheme $(\widetilde{A}, \widetilde{W})$ engages in sorting (pooling) are consistent with the intervals on which (A, W) engages in sorting (pooling), and that (A, W) and $(\widetilde{A}, \widetilde{W})$ jump at the same points.

   First, suppose that (A, W) entails sorting on an interval $[\theta_1, \theta_2] \subset [\underline{\theta}, \overline{\theta}]$. Using (9), the interim profit for $\theta \in [\theta_1, \theta_2]$ is then given by

$$R(\theta, \theta; A) - A(\theta) - W(\theta) = R(\theta_2, \theta_2; A) - A(\theta_2) - W(\theta_2) + \int_\theta^{\theta_2} D(p(x)) M(x; A) dx, \qquad (10)$$

where $M(x; A) = \frac{U}{N} + [1 - F(x)]^{N-1} I$. This equation can be rewritten as

$$A(\theta) + W(\theta) - [A(\theta_2) + W(\theta_2)] = - \int_\theta^{\theta_2} r(p(x), x)[\partial M(x; A)/\partial x] dx.$$

As shown in the proof of Proposition 1, the RHS equals the difference between type θ's and type θ_2's advertising levels in the advertising (Nash) equilibrium. Thus, $\widetilde{A}(\theta) \equiv A(\theta) + W(\theta)$ also entails sorting on the interval $[\theta_1, \theta_2]$ and satisfies (On-IC). Second, suppose that (A, W) entails pooling on an interval $[\theta_1, \theta_2] \subset [\underline{\theta}, \overline{\theta}]$. The interim profit for $\theta \in [\theta_1, \theta_2]$ takes the same form as in (10).

24

When market shares are constant on the pooling interval, then (10) can be rewritten as

$$A(\theta) + W(\theta) = A(\theta_2) + W(\theta_2), \tag{11}$$

and thus $\widetilde{A}(\theta) \equiv A(\theta) + W(\theta)$ also entails pooling on the interval $[\theta_1, \theta_2]$ and satisfies (On-IC). Since advertising is constant on the pooling interval, wars also are constant by (11). Thus, over a pooling interval, (On-IC) is maintained when a scheme with a constant level of current advertising and a future advertising war is replaced with a scheme in which the constant level of current advertising is raised and the possibility of a future advertising war is removed. Third, suppose that (A, W) involves a jump of market-share allocation at a point $\theta^* \in [\underline{\theta}, \overline{\theta}]$ such that

$$M(\theta^*; A) > \limsup_{\theta > \theta^*} M(\theta; A) \equiv M_+(\theta^*; A).$$

The associated limit for wars and advertising are denoted by $W_+(\theta^*)$ and $A_+(\theta^*)$, respectively. Incentive compatibility at the point θ^* implies that

$$A(\theta^*) + W(\theta^*) - [A_+(\theta^*) + W_+(\theta^*)] = r(p(\theta^*), \theta^*) \left[M(\theta^*; A) - M_+(\theta^*; A) \right].$$

Thus $\widetilde{A}(\theta) \equiv A(\theta) + W(\theta)$ entails a jump at θ^* and satisfies (On-IC).[28] Note lastly that when $(\widetilde{A}, \widetilde{W} \equiv 0)$ preserves the initial market-share allocation, $\widetilde{A}(\theta) \equiv A(\theta) + W(\theta)$ also preserves the interim profit. ∎

Lemma 2 identifies a substitutability between current advertising expenditures and future advertising wars. When a scheme (A, W) sets $W > 0$ for some values of θ and satisfies (On-IC), we understand that the expected future payoff is reduced due to the possibility of an advertising war. Lemma 2 indicates that we may then construct a point-wise equivalent scheme $(\widetilde{A}, \widetilde{W})$, in which the possibility of a future advertising war is eliminated $(\widetilde{W} \equiv 0)$ and current advertising expenditures are increased accordingly $(\widetilde{A}(\theta) \equiv A(\theta) + W(\theta))$. Wars are in this sense redundant.

Together, Lemma 2 and Proposition 6 greatly simplify our analysis. According to Lemma 2, for any (A, W) that solves the Relaxed Program, there exists an equivalent no-wars scheme, $(A^*, W^* \equiv 0)$ with $A^*(\theta) \equiv A(\theta) + W(\theta)$, that also solves the Relaxed Program. By Proposition 6, if firms are sufficiently patient, we may conclude that an optimal SPPE exists that is stationary and in which firms use A^* after all equilibrium path histories.[29] Proposition 6 also provides a program that may be solved in order to characterize A^*.

Our next step is to write the program identified in Proposition 6 in a more useful form. In

_____

[28] If (A, W) satisfies (On-IC), then $M(\theta; A)$ must be nonincreasing. As no type would "pay" more for less market share, incentive compatibility thus requires that $A(\theta) + W(\theta)$ is nonincreasing as well. It follows that $\widetilde{A}(\theta) \equiv A(\theta) + W(\theta)$ is nonincreasing.

[29] The arguments developed here may also be applied to sets of SPPE. For example, consider the set of SPPE in which full sorting occurs in each period. Proposition 6 also holds with respect to this class of equilibria; thus, using Lemma 2, we may conclude that an optimal SPPE within the full sorting class is the stationary SPPE in which firms use the advertising equilibrium of the stage game in every period. Thus, for firms to improve on the Nash equilibrium of the stage game, they must use an advertising scheme that entails some pooling.

particular, using Lemma 1 and $W^* \equiv 0$, we may integrate by parts and rewrite the program that A^* must solve as follows:

No-Wars Program: The program chooses A to maximize

$$E_\theta \left[R(\theta, \theta; A) - A(\theta) \right] = R(\bar{\theta}, \bar{\theta}; A) - A(\bar{\theta}) + E_\theta \left[D(p(\theta)) \frac{F}{f}(\theta) M(\theta; A) \right] \qquad (12)$$

subject to: $A(\theta)$ is nonincreasing in θ.

We may now summarize our discussion with the following proposition.

Proposition 7. *Let A^* solve the No-Wars Program. Then, there exists $\widehat{\delta} \in (0,1)$ such that, for all $\delta \geq \widehat{\delta}$, there exists an optimal SPPE which is stationary, wherein firms use A^* after all equilibrium-path histories.*

We emphasize that our no-wars (stationarity) finding is quite general, in that it holds for any demand function D and also for any distribution function F.

4.2 Optimal SPPE: Pooling at Zero Advertising

We now characterize A^* and thereby an optimal SPPE for patient firms that is stationary. To this end, we solve the No-Wars Program. We thus characterize the nonincreasing advertising scheme that maximizes expected profit, where as (12) indicates expected profit is comprised of profit at the top and expected information rents.

We encounter a related problem in Proposition 2, where we provide conditions under which expected profit is higher in the random equilibrium than in the advertising equilibrium. Generalizing beyond that particular comparison, we now show that the same conditions ensure that pooling at zero advertising in fact solves the No-Wars Program.

Proposition 8. *For δ sufficiently high, if F is log-concave and demand is sufficiently inelastic, or if the support of possible cost types is sufficiently small, then there exists an optimal SPPE that is stationary, wherein firms pool with zero advertising following all equilibrium-path histories.*

Proof. Using Proposition 7, we must show $A^p \equiv 0$ solves the No-Wars Program, if F is log-concave and demand is sufficiently inelastic or if $\bar{\theta} - \underline{\theta}$ is sufficiently small. Let A denote any other nonincreasing scheme. Note that $M(\theta; A)$ is then also nonincreasing, and recall that $M(\theta; A^p) \equiv \frac{1}{N}$. Consider first the profit at the top term in (12). If A entails any sorting, then $M(\bar{\theta}; A^p) = \frac{1}{N} > M(\bar{\theta}; A)$ and $A^p(\bar{\theta}) = 0 \leq A(\bar{\theta})$. Alternatively, if A is a pooling scheme (at some positive level of advertising), then $M(\bar{\theta}; A^p) = \frac{1}{N} = M(\bar{\theta}; A)$ and $A^p(\bar{\theta}) = 0 < A(\bar{\theta})$. In either case, profit at the top is strictly higher under A^p than A. Consider second the expected information rents term in (12). Note that expected information rents converge to zero as $\bar{\theta} - \underline{\theta}$ approaches zero; thus, the profit-at-the-top term dominates if the support of possible cost types is sufficiently small. Allowing for the general case in which the support may be large, we may define distribution functions $G(\theta; A)$ and $G(\theta; A^p)$ as in the proof of Proposition 2. Notice that $M(\theta; A^p) \equiv \frac{1}{N}$ crosses

$M(\theta; A)$ from below.[30] As in the proof of Proposition 2, if $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing, then expected information rents are weakly higher under A^p than A. Thus, when F is log-concave and demand is sufficiently inelastic, expected information rents are weakly higher under A^p than A. ∎

This result establishes conditions under which an optimal SPPE exists and entails pooling at zero advertising in all periods. The result thereby provides a formal confirmation of the idea that, even if advertising is legal and informed consumers are responsive to it, firms can still eliminate advertising as part of an optimal self-enforcing collusive agreement. When firms collude in this way, the welfare of informed consumers is reduced from the welfare that they enjoy in the non-cooperative advertising equilibrium. This is because the collusive agreement prevents informed consumers from using advertising to locate the lowest price in the market.

While profit at the top is uniquely maximized when firms pool at zero advertising, the maximization of expected information rents involves conflicting considerations. When the distribution function is log-concave, a pooling scheme shifts market share to higher-cost types and is attractive for this reason; however, when demand is downward sloping, a separating scheme is attractive, since it shifts market share to lower-cost types. Allowing for a wide range of advertising schemes, Proposition 8 isolates conditions under which the forces in favor of pooling at zero advertising dominate. In particular, if the distribution function is log–concave and demand is sufficiently inelastic, so that $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing, then pooling at zero advertising is an optimal SPPE for patient firms. While inelastic demand is sufficient in this sense, it is not necessary. As noted in Section 2, for a constant elasticity demand function such that $\epsilon > 1$, if types are distributed uniformly and $\overline{\theta}/[\overline{\theta} - \underline{\theta}] > \epsilon$, then $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing.

4.3 Optimal SPPE: Partial Pooling

While Proposition 8 characterizes an optimal SPPE under an important set of conditions, it is also interesting to consider optimal SPPE when these conditions fail. In this subsection, we maintain the assumption that firms are sufficiently patient, so that off-schedule constraints may be ignored, and show that under general conditions an optimal SPPE involves at least partial pooling.

A difficulty with solving the No-Wars Program is that the market-share function and the associated expected profit are conditional on the entire advertising schedule. Our analysis therefore proceeds from the fact that the entire advertising schedule can be decomposed into three different kinds of components: sorting, pooling and jumps. Consider the simplest case that has three parts: from the lowest step (from the highest type), a schedule has a pooling interval with $A(\theta) = 0$ on $(y, \overline{\theta}]$ and then jumps to a sorting interval on $[\underline{\theta}, y]$. This nondecreasing scheme has the following

---

[30]If A is a pooling scheme, then $M(\theta; A^p)$ crosses $M(\theta; A)$ from below in a weak sense.

expected profit:[31]

$$E_\theta[R(\theta,\theta;A) - A(\theta)] = r(p(\overline{\theta}),\overline{\theta})M(y,\overline{\theta};A) + \int_{\underline{\theta}}^{y} D(p(x))\frac{F}{f}(x)M(\underline{\theta},y;A)f(x)dx \quad (13)$$

$$+ \int_{y}^{\overline{\theta}} D(p(x))\frac{F}{f}(x)M(y,\overline{\theta};A)f(x)dx,$$

where (with some abuse in notation) $M(\underline{\theta},y;A)$ represents the market share for type x in the sorting interval $[\underline{\theta},y]$ and $M(y,\overline{\theta};A)$ represents the (constant) market share for any type in the pooling interval $(y,\overline{\theta}]$. Formally, these market share functions are defined as:

$$M(\underline{\theta},y;A) \equiv \frac{U}{N} + [1 - F(x)]^{N-1}I \text{ and } M(y,\overline{\theta};A) \equiv \frac{U}{N} + [1 - F(y)]^{N-1}\frac{I}{N}.$$

The level of jump is determined such that incentive-compatibility constraint is binding at y:

$$A(y) = r(p(y),y)\left[M(\underline{\theta},y;A) - M(y,\overline{\theta};A)\right],$$

where $M(\underline{\theta},y;A)$ is here evaluated at $x = y$. Note that when $y \to \overline{\theta}$, the scheme approaches the fully sorting scheme. Given the assumption that $p(\overline{\theta}) > \overline{\theta}$ and $f(\overline{\theta}) > 0$, we may differentiate (13) with respect to y and confirm that fully sorting can be improved upon by a scheme that has a pooling interval $(y,\overline{\theta}]$ at the top.

This finding illustrates a more general point. As we establish next, any incentive compatible scheme that has a sorting interval at the top can be improved upon by an alternative scheme that has a pooling interval at the top.

Lemma 3. *For any F, if δ is sufficiently high, then any optimal SPPE that is stationary has a pooling interval $(y,\overline{\theta}]$ on which $A(\theta) = 0$.*

The proof is in the Appendix. Since the repeated play of the advertising equilibrium of the static game is a stationary SPPE that entails full sorting, Lemma 3 ensures that for sufficiently patient firms an optimal SPPE must involve some pooling and strictly improve upon the repeated use of the advertising equilibrium.

Lemma 3 thus establishes that under general conditions an optimal SPPE involves at least partial pooling. In the supplementary materials for this paper, we extend the analysis and offer a more complete characterization of optimal SPPE for any F when δ is sufficiently high. In particular, we show that an optimal SPPE entails a single flat step on any interval where $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing. We show as well that, if an optimal SPPE entails sorting over an interval, then $D(p(\theta))\frac{F}{f}(\theta)$ must be decreasing over that interval.

---

[31]The expected profit is derived in the Appendix.

4.4 Off-Schedule Incentive Constraints

Up to this point, we have ignored the off-schedule constraints by assuming that firms are sufficiently patient. We now consider off-schedule constraints and characterize the critical discount factor, $\widehat{\delta} \in (0,1)$, above which an optimal SPPE exists that is stationary (as established in Proposition 7). To focus our discussion, we emphasize the setting described in Proposition 8, wherein the optimal SPPE is stationary and entails pooling at zero advertising following all equilibrium-path histories.

Suppose, then, that the optimal SPPE entails pooling at zero advertising. When firms behave in this fashion, a firm faces a temptation to cheat by advertising a small, positive amount, as it thereby attracts all informed consumers rather than only its share of these consumers. This short-term incentive to cheat must be balanced against the long-term cost of a punishment (i.e., a reduced continuation value). Given our focus on SPPE, such a punishment must be experienced by all firms. We thus suppose that an off-schedule deviation of this kind triggers a reversion to the advertising equilibrium of the static game.[32] Thus, the long-term cost of an off-schedule deviation is that the future discounted expected profit associated with pooling at zero advertising is replaced with that associated with the repeated play of the advertising equilibrium. In other words, if a firm cheats on the collusive agreement to not advertise, then a breakdown in cooperation occurs and the firms revert to the advertising equilibrium thereafter.

We now consider the type of firm for which the off-schedule constraint first binds. Given our assumption that cost types are determined in an iid fashion through time, a firm faces the same long-term cost of an off-schedule deviation regardless of its current type, θ. The short-term incentive to deviate, however, is sensitive to θ. In particular, when firms pool at zero advertising, a firm with cost type $\underline{\theta}$ has the greatest short-term incentive to defect. This type of firm values most the increase in market share that accompanies cheating, since it has the highest profit-if-win, $r(p(\underline{\theta}), \underline{\theta})$. When firms pool at zero advertising, the off-schedule constraint is sure to hold for all θ if it holds for $\underline{\theta}$. We may thus represent the off-schedule constraint for this situation as follows:

$$r(p(\underline{\theta}), \underline{\theta})I(1 - \frac{1}{N}) \leq \frac{\delta}{1-\delta}[\pi^p - \pi^s], \tag{14}$$

where $\pi^s \equiv E_\theta [\Pi(\theta, \theta; A)]$ and $\pi^p \equiv E_\theta \left[r(p(\theta), \theta)\frac{1}{N} \right]$ are a firm's expected per-period profit when firms separate using the advertising equilibrium, A, and pool at zero advertising, respectively. These profit terms are formally characterized in (3) and (4).

Solving (14) for the critical discount factor, we obtain that pooling at zero advertising satisfies the off-schedule constraint if

$$\delta \geq \widehat{\delta}^p \equiv \frac{r(p(\underline{\theta}), \underline{\theta})(N-1)I}{r(p(\underline{\theta}), \underline{\theta})(N-1)I + N(\pi^p - \pi^s)}.$$

As shown in Proposition 2, $\pi^p > \pi^s$ if F is log-concave and demand is sufficiently inelastic or if

---

[32]Other symmetric punishments, such as those that take a "carrot-stick" form, may also be considered. Building on arguments developed by Athey, Bagwell and Sanchirico (2004), we can show that the repeated play of the advertising equilibrium generates the lowest SPPE payoff when $D(p(\theta))\frac{F}{f}(\theta)$ is everywhere nondecreasing.

$\overline{\theta} - \underline{\theta}$ is sufficiently small. Thus, under these conditions, $\widehat{\delta}^p \in (0,1)$. We have thus established:

Proposition 9. *If F is log-concave and demand is sufficiently inelastic, or if the support of possible cost types is sufficiently small, then $\widehat{\delta}^p \in (0,1)$ and, for all $\delta \geq \widehat{\delta}^p$, there exists an optimal SPPE which is stationary, wherein firms pool at zero advertising after all equilibrium-path histories.*

In comparison to Proposition 7, Proposition 9 provides an explicit characterization of the critical discount factor above which firms can enforce an optimal SPPE in which they agree to eliminate the use of advertising.

It is also possible to derive characterizations of optimal collusion when the discount factor is not sufficiently high to support the scheme that would be optimal were only on-schedule constraints considered. We can show, for example, that the no-wars finding extends to the low-δ setting. Intuitively, if the off-schedule constraint is an issue, it is better to shift current-period profit toward the future, as a firm then has more to lose in the future by undertaking an off-schedule deviation in the present. Exploiting the substitutability between current advertising and future wars, firms can achieve the desired shift by increasing advertising and eliminating future wars. Athey, Bagwell and Sanchirico (2004) provide a related argument in their analysis of price collusion, and so we do not develop this point in detail here.

5 Sequential Search

We assume above that consumers are unable to engage in sequential search. Focusing on the static setting, we now examine equilibrium behavior when this assumption is relaxed. Thus, we allow that after a consumer visits a firm and observes that firm's price, the consumer may elect to incur a search cost and visit another firm.

Consider then a modified static game, in which consumers can undertake costly sequential search and firms choose advertising levels and prices. A Symmetric Perfect Bayesian Equilibrium may be informally defined in terms of the following requirements: (i) each firm selects its advertising level and price to maximize its expected profit, given its type and the strategies of other players; (ii) each consumer selects an initial firm to visit and any subsequent firm to visit in a way that maximizes the consumer's expected welfare at each point, given the information that the consumer then has and the consumer's beliefs about prices at firms not yet visited; (iii) where possible, consumers' beliefs are formed in a manner consistent with Bayes' rule, given the equilibrium strategies of firms;[33] and (iv) firms use symmetric price and advertising strategies. An advertising equilibrium is a Symmetric Perfect Bayesian Equilibrium in which informed consumers pick an initial firm using the advertising search rule while uninformed consumers pick an initial firm at random. A random

---

[33]The concept of Perfect Bayesian Equilibrium also includes a no-signaling-what-you-don't-know requirement. In the present context, this means that, if a consumer initially visits firm i and contemplates undertaking the sequential search cost and visiting some other firm j, then the consumer's belief about the price that might be observed at firm j is not altered by the price observed at firm i. Of course, for an informed consumer, the belief about the price at firm j may be influenced by the advertising level selected by firm j.

equilibrium is a Symmetric Perfect Bayesian Equilibrium in which all consumers ignore advertising and select an initial firm at random.

We begin by observing that the sequential-search option is irrelevant if the cost of sequential search is sufficiently large relative to the expected dispersion of prices in the market. Suppose that firms follow the advertising equilibrium of the original static game as characterized in Proposition 1. An uninformed consumer is then most tempted to search again in the event that the consumer encounters the highest possible monopoly price, $p(\overline{\theta})$. Let $U(p)$ denote consumer surplus at the price p, and let the cost of sequential search be denoted as $d > 0$.[34] Even a consumer that encounters $p(\overline{\theta})$ won't gain from sequential search, if $U(p(\overline{\theta})) \geq E_\theta U(p(\theta)) - d$. Thus, if $p(\overline{\theta}) - E_\theta p(\theta)$ is small relative to the cost of sequential search, then an uninformed consumer never gains from sequential search. This condition is sure to hold in the limiting case of perfectly inelastic demand, since then the monopoly price is independent of production costs. Likewise, for any CES demand with elasticity $\epsilon > 1$, we have that $p(\overline{\theta}) - E_\theta p(\theta) = \frac{\epsilon}{\epsilon-1}[\overline{\theta} - E\theta]$. Thus, if the extent of dispersion in production costs is small relative to the size of the sequential-search cost, then uniformed consumers will not search again even after encountering the highest monopoly price.

If instead the cost of sequential search is small relative to the expected dispersion of prices, then higher-cost firms induce search if they select their monopoly prices. To capture this situation, we assume henceforth that $U(p(\overline{\theta})) < E_\theta U(p(\theta)) - d$. Building on work by Reinganum (1979) and Bagwell and Ramey (1996), our goal is to establish conditions under which an advertising equilibrium exists in which firms with cost types at or above a critical level $\theta_c \in (\underline{\theta}, \overline{\theta})$ select the monopoly price for this critical type. In particular, we seek to construct an advertising equilibrium in which a firm with cost type $\theta > \theta_c$ prices at $p(\theta_c) < p(\theta)$, where $p(\theta_c)$ is determined so that the costs and benefits of sequential search are equal. Higher-cost firms then "limit price" and thereby deter uninformed consumers from searching again.

In our proposed advertising equilibrium, a firm of cost type θ thus selects the price $p^*(\theta) \equiv \min\{p(\theta), p(\theta_c)\}$ and earns the corresponding net revenue $r(p^*(\theta), \theta)$. We now impose a new assumption that $p(\underline{\theta}) > \overline{\theta}$. This assumption is sure to hold if the dispersion in cost types is not too great or if demand is sufficiently inelastic, and it ensures that $p(\theta_c) > \overline{\theta}$ so that $r(p^*(\theta), \theta)$ remains strictly positive even for the highest type. Observe also that $r(p^*(\theta), \theta)$ is strictly decreasing with $\frac{dr(p^*(\theta), \theta)}{d\theta} = -D(p^*(\theta)) < 0$. With these properties in place, we can confirm that the arguments used in the proof of Proposition 1 continue to hold when firms use the pricing function $p^*(\theta)$. Thus, the level of advertising again strictly declines as costs increase, and no firm of any type gains from undertaking an on-schedule deviation and mimicking the advertising level of some other type. Informed consumers are again rational in visiting the firm with the highest advertising level, since this firm selects the lowest price in the market.[35] Two issues remain. First, we must establish that a critical value $\theta_c \in (\underline{\theta}, \overline{\theta})$ indeed exists such that an uninformed consumer is indifferent to sequential search upon observing $p(\theta_c)$. Second, we must establish that no firm with cost type $\theta > \theta_c$ would

---

[34] For simplicity, we assume that the initial search has zero cost.

[35] Note, though, that informed consumers are indifferent about using the advertising search rule in the event that all firms draw cost types at or above θ_c.

gain from undertaking an off-schedule deviation to a higher price.

Consider the first issue. Under our assumption that $U(p(\overline{\theta})) < E_\theta U(p(\theta)) - d$, it is straightforward to establish that there exists a unique value $\theta_c \in (\underline{\theta}, \overline{\theta})$ such that

$$U(p(\theta_c)) = [1 - F(\theta_c)]U(p(\theta_c)) + \int_{\underline{\theta}}^{\theta_c} U(p(\theta))dF(\theta) - d. \qquad (15)$$

The LHS of (15) represents the consumer welfare from remaining with a firm that selects $p(\theta_c)$, while the RHS represents the expected welfare from incurring the sequential-search cost d and finding the same price or a lower price. The critical value $\theta_c \in (\underline{\theta}, \overline{\theta})$ is then determined so as to make the consumer indifferent between the two options. Notice that θ_c is independent of the fraction of informed consumers, I, and is strictly increasing in the sequential-search cost, d. As d gets close to zero, θ_c gets close to $\underline{\theta}$ and thus almost all types select the limit price.

To understand the second issue, consider a firm with cost type $\theta > \theta_c$. This firm retains its uninformed consumers if it sets the limit price, $p(\theta_c)$, and loses its uninformed consumers if it sets any higher price. Under our assumption that $p(\underline{\theta}) > \overline{\theta}$, we know that the firm earns strictly positive net revenue on its uninformed consumers at the price $p(\theta_c)$. Thus, as regards its uninformed consumers, the firm earns strictly more by selecting the price $p(\theta_c)$ than it would make by undertaking an off-schedule deviation to any higher price. But this firm must also consider informed consumers. With probability $[1 - F(\theta)]^{N-1}$, this firm advertises more than all other firms and receives the informed consumers. In this event, as in the model analyzed by Bagwell and Ramey (1996), the informed consumers observe all advertising choices and thus know that all other firms have higher costs and thus select the price $p(\theta_c)$. The informed consumers will then tolerate a price hike without searching again, provided that the hike is not too large. The maximal price hike that informed consumers will tolerate is $h(d)$ where $h(d)$ is defined by $U(p(\theta_c) + h(d)) \equiv U(p(\theta_c)) - d$. It follows that the optimal off-schedule deviation for a firm of type $\theta > \theta_c$ is the price $p(\theta, \theta_c, d) \equiv \min\{p(\theta), p(\theta_c) + h(d)\}$, where θ_c is determined as a function of d by (15).

We may now conclude that a firm with cost type $\theta > \theta_c$ does not gain from an off-schedule deviation to a higher price if

$$\Omega(\theta, \theta_c, d) \equiv [(1 - F(\theta))^{N-1}I][r(p(\theta, \theta_c, d), \theta) - r(p(\theta_c), \theta)] - \frac{U}{N}r(p(\theta_c), \theta) \leq 0. \qquad (16)$$

The first term on the RHS of (16) captures the possible benefit of a price hike in terms of more profitable sales to informed consumers whereas the second term reflects the certain cost of a price hike in terms of lost sales to uninformed consumers. Notice that $\Omega(\theta_c, \theta_c, d) < 0$, since $p(\theta_c, \theta_c, d) = p(\theta_c)$. Likewise, $\Omega(\overline{\theta}, \theta_c, d) < 0$ follows, since the highest-cost firm wins the informed consumers with probability zero and earns strictly positive net revenue at the price $p(\theta_c)$ under our assumption that $p(\underline{\theta}) > \overline{\theta}$. Outside of these boundary cases, we cannot immediately sign $\Omega(\theta, \theta_c, d)$. We can, however, state the following sufficient condition: There exists $I^* \in (0, 1)$ such that if $I < I^*$ then for all $\theta \in (\theta_c, \overline{\theta})$, $\Omega(\theta, \theta_c, d) < 0$. In other words, if the fraction of informed consumers is not too

great, then no type of firm will undertake an off-schedule deviation by raising price.

We may now summarize our findings as follows.

Proposition 10. *Consider the static game, modified to allow for sequential search. Assume that the search cost satisfies $U(p(\overline{\theta})) < E_\theta U(p(\theta)) - d$ and that $p(\underline{\theta}) > \overline{\theta}$. There exists $I^* \in (0,1)$ such that if $I < I^*$ then an advertising equilibrium exists. In this equilibrium, (i) firms use an advertising strategy $A(\theta)$ that is strictly decreasing and differentiable and satisfies $A(\overline{\theta}) = 0$; (ii) firms use the pricing strategy $p^*(\theta)$, where $\theta_c \in (\underline{\theta}, \overline{\theta})$ satisfies (15); and (iii) consumers do not engage in sequential search along the equilibrium path.*

In effect, Proposition 10 establishes conditions under which Proposition 1 extends to the setting in which sequential search is possible and not prohibitively expensive.[36]

We now consider the effect of sequential search on the comparison between expected profits under the random and advertising equilibria. When sequential search is possible, our assumption that $p(\underline{\theta}) > \overline{\theta}$ ensures that a random equilibrium exists, wherein firms use the modified pricing schedule, $p^*(\theta)$.[37] As this assumption implies that profit at the top is strictly positive, the random equilibrium again generates strictly greater profit at the top than does the constructed advertising equilibrium (when it exists). When sequential search is prohibited, expected information rents are higher under the random than advertising equilibrium if $\frac{F}{f}(\theta)D(p(\theta))$ is nondecreasing. Likewise, when sequential search is possible, expected information rents are strictly higher under the random than advertising equilibrium if $\frac{F}{f}(\theta)D(p^*(\theta))$ is nondecreasing. Since $p^*(\theta)$ is constant in θ for $\theta > \theta_c$, log-concavity of F alone now ensures that $\frac{F}{f}(\theta)D(p^*(\theta))$ is nondecreasing when $\theta > \theta_c$. Thus, the tension between log-concavity and reduced demand is removed for higher types when sequential search is possible. In this respect, the possibility of sequential search serves to *strengthen* our basic result that firms achieve higher expected profit when they restrict the use of advertising.[38]

6 Other Extensions

In this section, we provide an informal discussion of extended models in which advertising enters the demand function and asymmetric PPE are allowed.

6.1 Advertising in the Demand Function

In our analysis above, advertising does not enter the demand function. Instead, we build on earlier work and study the role of advertising in directing informed consumers to the lowest prices. Of

---

[36]The advertising equilibrium of the modified static game is also unique, if the definitions of the advertising and random search rules are extended to cover sequential search decisions. Otherwise, some uninformed consumers that encounter the price $p(\theta_c)$ may undertake sequential search out of indifference, for example.

[37]The existence of the random equilibrium does not require any additional assumption on the fraction of informed consumers, since firms do not advertise in the random equilibrium and thus all consumers are, in effect, uninformed. Thus, the random equilibrium is the counterpart of the equilibrium featured by Reinganum (1979).

[38]Note, though, that sequential search lowers profit at the top, since higher-cost firms earn lower profit when sequential search is possible. Sequential search thus diminishes the magnitude of the profit-at-the-top advantage that the random equilibrium has in comparison to the advertising equilibrium.

course, other kinds of advertising are also interesting and worthy of separate analysis. Here, we briefly highlight two possibilities.

A first possibility is that advertising by a firm directly increases the demand function that informed consumers bring to this firm, in the event that the firm out-advertises all other firms. In this case, Proposition 4 may change in interesting ways, as it is no longer clear that firms earn greater expected profit in the pricing equilibrium than in the advertising equilibrium. As noted, the low prices of the pricing equilibrium serve to expand in-store demand, and they thus elevate expected information rents. If advertising enters the demand function directly, then high advertising likewise expands in-store demand and thereby elevates expected information rents. Whether price or non-price advertising is more profitable may then depend on the respective elasticities of demand with respect to price and advertising.

A second possibility is that advertising by any one firm may have a public-good flavor and serve to expand the size of market demand. By contrast, in the model analyzed above, advertising is redistributive: the size of aggregate demand is not affected by advertising, and so one firm's market-share gain is another firms' market-share loss. In the case of public-good advertising, when a firm advertises more, aggregate demand increases and so rival firms benefit to some degree as well. Such advertising may be especially important for new-product markets. An analysis of this kind of advertising is an important direction for future work.

6.2 Asymmetric PPE

As Athey and Bagwell (2001) show, when firms collude in prices, profit may be higher in asymmetric PPE than in SPPE. They emphasize the role of future market share favors, whereby a firm that claims low costs and enjoys high market share today must suffer a reduced market share in the future. Rival firms then enjoy a future market share gain. Thus, asymmetric PPE allow that continuation values may be used to satisfy on-schedule constraints, without requiring that all firms symmetrically experience a reduced continuation value.[39]

In the price-collusion model, consumers directly observe price and have no independent interest in firms' costs. By contrast, in the advertising model analyzed in this paper, informed consumers observe advertising and draw inferences as to costs and thus prices. The construction of asymmetric PPE may be more challenging in this context. Suppose, for example, that one firm advertises heavily in the current period and that the equilibrium then requires that this firm advertise less in the future, so as to transfer future market share to other firms. Consider now the informed consumers. If they understand the equilibrium, then they recognize that the reduced level of advertising by this firm in some future period is not necessarily a signal that this firm has a high cost type and thus a high price in that period. Thus, even if the equilibrium calls for reduced advertising by this firm, this in itself does not guarantee that the firm obtains reduced market share.

---

[39] As Athey and Bagwell (forthcoming) show in their analysis of price collusion, however, when cost shocks are persistent, the advantage of asymmetric PPE may be significantly reduced. Indeed, if demand is perfectly inelastic and the distribution of types is log-concave, they show that a stationary pooling equilibrium is optimal for patient firms when cost types are perfectly persistent.

7 Conclusion

We investigate the advertising behavior of firms with private information as to their respective costs. We first analyze a static model, and we show there that an advertising equilibrium exists, in which informed consumers use an advertising search rule whereby they buy from the highest-advertising firm. The key point is that the highest-advertising firm has the lowest cost and thus selects the lowest price. In this way, "non-informative" advertising directs consumers to the lowest price in the market. We establish conditions under which firms earn greater expected profit when advertising is banned. Consumer welfare falls in this case, however. We then analyze a dynamic model in which privately informed firms interact repeatedly. In this setting, firms may achieve a collusive equilibrium in which they limit the use of advertising, and we establish conditions under which optimal collusion entails pooling at zero advertising. In summary, advertising can promote product efficiency and raise consumer welfare; however, firms often have incentive to diminish advertising competition, whether through regulatory restrictions or collusion.

8 Appendix

Proof of Proposition 5. (ii) Note first that advertising at the top is held fixed at $A(\overline{\theta}) = A'(\overline{\theta}) = 0$ for all N. Differentiating $|A'(\theta)|$ with respect to N yields:

$$\frac{\partial |A'(\theta)|}{\partial N} = |A'(\theta)| \left(\frac{1 + (N-1)\ln[1 - F(\theta)]}{N-1} \right).$$

The equation means that for a slight increase of N, $A(\theta)$ becomes flatter over the types above $F^{-1}(1 - e^{-\frac{1}{N-1}}) \in (\underline{\theta}, \overline{\theta})$ and steeper over the types below $F^{-1}(1 - e^{-\frac{1}{N-1}})$. We can next show that advertising at the bottom, $A(\underline{\theta})$, strictly increases when N rises. To see this, integrating by parts, we get

$$
\begin{aligned}
A(\underline{\theta}) &= \int_{\underline{\theta}}^{\overline{\theta}} r(p(x), x)(N-1)[1 - F(x)]^{N-2} f(x) I dx \\
&= r(p(\underline{\theta}), \underline{\theta}) I - \int_{\underline{\theta}}^{\overline{\theta}} [1 - F(x)]^{N-1} D(p(x)) I dx.
\end{aligned}
$$

The integral on the RHS strictly decreases with N and thus $A(\underline{\theta})$ strictly increases in N. Hence, we can now conclude that there exists a cutoff type $\widehat{\theta} < F^{-1}(1 - e^{-\frac{1}{N-1}})$ such that equilibrium advertising strictly increases with N for $\theta \in [\underline{\theta}, \widehat{\theta})$, strictly decreases with N for $\theta \in (\widehat{\theta}, \overline{\theta})$, and is constant with N when $\theta \in \{\widehat{\theta}, \overline{\theta}\}$.

(iv) For the proof, we proceed with four steps as follows. First, we establish a monotonicity in the ratio of advertising equilibrium slopes under MLR dominance. Define

$$\gamma(\theta) \equiv \frac{A'_F(\theta)}{A'_G(\theta)} = \frac{f(\theta)}{g(\theta)} \left[\frac{1 - F(\theta)}{1 - G(\theta)} \right]^{N-2}.$$

For $\theta \in [\underline{\theta}, \overline{\theta})$, the ratio $\gamma(\theta)$ of two slopes is strictly increasing in θ, since $\frac{f(\theta)}{g(\theta)}$ and $\frac{1-F(\theta)}{1-G(\theta)}$ are then positive and strictly increasing under MLR dominance. The latter term, $\frac{1-F(\theta)}{1-G(\theta)}$, is strictly increasing if $\frac{1-F(\theta)}{f(\theta)} > \frac{1-G(\theta)}{g(\theta)}$. To see that this inequality holds for $\theta \in [\underline{\theta}, \overline{\theta})$, note that MLR dominance can be re-stated as $\frac{f(y)}{f(x)} > \frac{g(y)}{g(x)}$ for all $y > x$; hence, for $x \in [\underline{\theta}, \overline{\theta})$, MLR dominance implies $\int_x^{\overline{\theta}} \frac{f(y)}{f(x)} dy > \int_x^{\overline{\theta}} \frac{g(y)}{g(x)} dy$ and thus $\frac{1-F(x)}{f(x)} > \frac{1-G(x)}{g(x)}$. Second, we establish that $A_F(\underline{\theta}) < A_G(\underline{\theta})$. Note that

$$A_F(\underline{\theta}) - A_G(\underline{\theta}) = -\int_{\underline{\theta}}^{\overline{\theta}} \left([1 - F(x)]^{N-1} - [1 - G(x)]^{N-1}\right) D(p(x)) I dx.$$

We thus have that $A_F(\underline{\theta}) < A_G(\underline{\theta})$ if $\frac{1-F(\theta)}{1-G(\theta)} > 1$ for all $\theta > \underline{\theta}$. This inequality holds, since $\frac{1-F(\theta)}{1-G(\theta)}$ achieves its minimum value of 1 at $\underline{\theta}$ and (as established above) is strictly increasing for $\theta \in [\underline{\theta}, \overline{\theta})$ under MLR dominance. Third, we show that $\gamma(\underline{\theta}) = \frac{f(\underline{\theta})}{g(\underline{\theta})} < 1 < [\frac{f(\overline{\theta})}{g(\overline{\theta})}]^{N-1} = \gamma(\overline{\theta})$. The stated properties for $\gamma(\underline{\theta})$ follow immediately from the definition of $\gamma(\theta)$ and MLR dominance, while the stated properties for $\gamma(\overline{\theta})$ follow from using L'Hopital's rule and MLR dominance. Given $A_F(\underline{\theta}) < A_G(\underline{\theta})$, $A_F(\overline{\theta}) = A_G(\overline{\theta}) = 0$ and $\gamma(\overline{\theta}) > 1$, we can conclude that there exists $\widetilde{\theta} \in (\underline{\theta}, \overline{\theta})$ at which $A_G(\theta)$ crosses $A_F(\theta)$ from above. Fourth, we establish that a second interior crossing does not exist. Assume to the contrary that there exists $\theta_2 \in (\underline{\theta}, \overline{\theta})$ at which $A_G(\theta)$ crosses $A_F(\theta)$ from below and thus $\gamma(\theta_2) > 1$. Given $A_F(\overline{\theta}) = A_G(\overline{\theta}) = 0$ and $\gamma(\overline{\theta}) > 1$, there must then exist $\theta_3 \in (\theta_2, \overline{\theta})$ at which $A_G(\theta)$ crosses $A_F(\theta)$ from above and thus $\gamma(\theta_3) < 1$. But this contradicts our first result that $\gamma(\theta)$ is strictly increasing in θ over $\theta \in [\underline{\theta}, \overline{\theta})$ under MLR dominance.

Derivation of Interim-Stage Profit. We show that if A has a pooling interval with $A(\theta) = 0$ on $(y, \overline{\theta}]$ and jumps to a sorting interval on $[\underline{\theta}, y]$, then it has the expected profit (13) in the text:

$$E_\theta[R(\theta, \theta; A) - A(\theta)] = r(p(\overline{\theta}), \overline{\theta}) M(y, \overline{\theta}; A) + \int_{\underline{\theta}}^y D(p(x)) \frac{F}{f}(x) M(\underline{\theta}, y; A) f(x) dx$$

$$+ \int_y^{\overline{\theta}} D(p(x)) \frac{F}{f}(x) M(y, \overline{\theta}; A) f(x) dx,$$

where $M(\underline{\theta}, y; A)$ and $M(y, \overline{\theta}; A)$ are defined as in the text. The interim-stage profit for $\theta \leq y$ is

$$R(\theta, \theta; A) - A(\theta) = R(y, y; A) - A(y) + \int_\theta^y D(p(x)) M(\underline{\theta}, y; A) dx,$$

while the interim-stage profit at y is

$$R(y, y; A) - A(y) = R(\overline{\theta}, \overline{\theta}; A) - A(\overline{\theta}) + \int_y^{\overline{\theta}} D(p(x)) M(y, \overline{\theta}; A) dx.$$

Using two equations, we find the interim-stage profit for $\theta \leq y$:

$$R(\theta, \theta; A) - A(\theta) = R(\overline{\theta}, \overline{\theta}; A) - A(\overline{\theta}) + \int_\theta^y D(p(x)) M(\underline{\theta}, y; A) dx + \int_y^{\overline{\theta}} D(p(x)) M(y, \overline{\theta}; A) dx.$$

The interim-stage profit for $\theta > y$ is

$$R(\theta, \theta; A) - A(\theta) = R(\overline{\theta}, \overline{\theta}; A) - A(\overline{\theta}) + \int_\theta^{\overline{\theta}} D(p(x))M(y, \overline{\theta}; A)dx.$$

Based on the two interim-stage profits, we find the expected value (13) in the text by integrating by parts and setting $A(\overline{\theta}) = 0$. ∎

Proof of Lemma 3. Suppose that a scheme has a sorting interval at the top on $(z, \overline{\theta}]$. Then we can consider an alternative scheme A that decomposes the sorting interval on $(z, \overline{\theta}]$ into a sorting interval on $(z, y]$ and a pooling interval on $(y, \overline{\theta}]$, where $y > z$. Letting $M(\theta; A)$ denote the market allocation for types below z, the expected profit becomes

$$
\begin{aligned}
E_\theta \left[R(\theta, \theta; A) - A(\theta) \right] &= r(p(\overline{\theta}), \overline{\theta}) \left[\frac{U}{N} + [1 - F(y)]^{N-1}\frac{I}{N} \right] + \int_{\underline{\theta}}^z D(p(x))F(x)M(x; A)dx \\
&\quad + \int_z^y D(p(x))F(x) \left[\frac{U}{N} + [1 - F(x)]^{N-1}I \right] dx \\
&\quad + \int_y^{\overline{\theta}} D(p(x))F(x) \left[\frac{U}{N} + [1 - F(y)]^{N-1}\frac{I}{N} \right] dx.
\end{aligned}
$$

If the initial advertising scheme is nonincreasing in θ, then the alternative scheme A also is nonincreasing in θ. Note that if $y \to \overline{\theta}$, then this scheme approaches the initial scheme. We show that the optimal choice of y is lower than $\overline{\theta}$. Taking derivatives of the objective function with respect to y, we obtain

$$
\begin{aligned}
\frac{\partial E_\theta \left[R(\theta, \theta; A) - A(\theta) \right]}{\partial y} &= \frac{N-1}{N}[1 - F(y)]^{N-1}I \left[D(p(y))F(y) - r(p(\overline{\theta}), \overline{\theta})\frac{f(y)}{1 - F(y)} \right] \\
&\quad - \int_y^{\overline{\theta}} D(p(x))F(x) \left[\frac{N-1}{N}[1 - F(y)]^{N-2}f(y)I \right] dx.
\end{aligned}
$$

Because of the assumption that $f(\overline{\theta}) > 0$ and $p(\overline{\theta}) > \overline{\theta}$, the expected profit rises when y slightly falls from $\overline{\theta}$. ∎

9 References

Abreu, D., D. Pearce and E. Stacchetti (1986), "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory*, 39.1, 251-69.

Abreu, D., D. Pearce and E. Stacchetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58.5, 1041-63.

Athey, S. (2002), "Monotone Comparative Statics under Uncertainty," *Quarterly Journal of Economics*, 117, 187-223.

Athey, S. and K. Bagwell (2001), "Optimal Collusion with Private Information," *Rand Journal of Economics*, 32.3, 428-65.

Athey, S. and K. Bagwell (forthcoming), "Collusion with Persistent Cost Shocks," *Econometrica*.

Athey, S., K. Bagwell and C. Sanchirico (2004), "Collusion and Price Rigidity," *Review of Economic Studies*, 71.2, 317-349.

Bagwell, K. (2007), "The Economic Analysis of Advertising," in M. Armstrong and R. Porter (eds.), **Handbook of Industrial Organization, Vol. 3**, North-Holland: Amsterdam, 1701-1844.

Bagwell, K. and G. Ramey (1988), "Advertising and Limit Pricing," *Rand Journal of Economics*, 19, 59-71.

Bagwell, K. and G. Ramey (1994a), "Coordination Economies, Advertising, and Search Behavior in Retail Markets," *American Economic Review*, 84.3, 498-517.

Bagwell, K. and G. Ramey (1994b), "Advertising and Coordination," *Review of Economic Studies*, 61, 153-171.

Bagwell, K. and G. Ramey (1996), "Coordination Economies, Sequential Search and Advertising," Northwestern CMSEMS D.P. No. 1148.

Bagwell, K. and A. Wolinsky (2002), "Game Theory and Industrial Organization," in R. J. Aumann and S. Hart (eds.), **Handbook of Game Theory, Vol. 3**, North-Holland: Amsterdam, 1851-1895.

Bagwell, K. and R. W. Staiger (2005), "Enforcement, Private Political Pressure and the GATT/WTO Escape Clause," *Journal of Legal Studies*, 34.2, 471-513.

Bain, J. (1956), **Barriers to New Competition: Their Character and Consequences in Manufacturing Industries**, Cambridge, MA: Harvard University Press.

Benham, L. (1972), "The Effect of Advertising on the Price of Eyeglasses," *Journal of Law and Economics*, 15.2, 337-352.

Cable, J. (1972), "Market Structure, Advertising Policy and Intermarket Differences in Advertising Intensity," in K. Cowling (ed.), **Market Structure and Corporate Behavior**, London: Gray-Mills Publishing, 107-24.

Cady, J. (1976), "An Estimate of the Price Effects of Restrictions on Drug Price Advertising," *Economic Inquiry*, 14.4, 493-510.

Comanor, W. S. and T. A. Wilson (1974), **Advertising and Market Power**, Cambridge, MA: Harvard University Press.

Demsetz, H. (1973), "Industry Structure, Market Rivalry, and Public Policy," *Journal of Law and Economics*, 16, 1-9.

Ferguson, J. M. (1974), **Advertising and Competition; Theory, Measurement, Fact**, Cambridge, MA.: Lippincott, Ballinger.

Friedman, J. W. (1983), "Advertising and Oligopolistic Equilibrium," *Bell Journal of Economics*, 14.2, 464-473.

Gasmi, F., Laffont, J.J., and Q. Vuong (1992), "Econometric Analysis of Collusive Behavior in a Soft-Drink Market," *Journal of Economics & Management Strategy*, 1.2, 277-311.

Greer, D. F. (1971), "Advertising and Market Concentration," *Southern Economic Journal*, 38.1, 19-32.

Hanazono, M. and H. Yang (2007), "Collusion, Fluctuating Demand and Price Rigidity," *International Economic Review*, 48.2, 483-515.

Hopkins, E., and T. Kornienko (2007) "Cross and Double Cross: Comparative Statics in First Price and All Pay Auctions," *The B. E. Journal of Theoretical Economics*, Vol. 7: Iss. 1 (Topics), Article 19.

Kadiyali, V. (1996), "Entry, Its Deterrence, and Its Accommodation: A Study of the US Photographic Film Industry," *RAND Journal of Economics*, 27.3., 452-478.

Lebrun, B. (1998), "Comparative Statics in First Price Auctions," *Games and Economic Behavior*, 25, 97-110.

Lee, G. M. (2007), "Trade Agreements with Domestic Policies as Disguised Protection," *Journal of International Economics*, 71.1, 241-259.

Maskin, E. and J. Riley (1984), "Optimal Auctions with Risk Averse Buyers," *Econometrica*, 52.6, 1473-1518.

McAfee, R.P. and J. McMillan (1992), "Bidding Rings," *American Economic Review*, 82.3, 579-599.

Milgrom, P. and I. Segal (2002), "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70.2, 583-601.

Nelson, P. (1974), "Advertising as Information," *Journal of Political Economy*, 82, 729-754.

Nocke, V. (2007), "Collusion and Dynamic (Under-) Investment in Product Quality," *Rand Journal of Economics*, 38.1, 227-49.

Reinganum, J. (1979), "A Simple Model of Equilibrium Price Dispersion," *Journal of Political Economy*, 87, 851-58.

Scherer, F. M. (1980), **Industrial Market Structure and Economic Performance**, 2d ed., Chicago: Rand McNally College Publishing Co.

Simon, J. (1970), **Issues in the Economics of Advertising**, Urbana: University of Illinois Press.

Spulber, D.F. (1995), "Bertrand Competition when Rivals' Costs are Unknown," *Journal of Industrial Economics*, 43.1, 1-11.

Stigler, G. J. (1968), "Price and Non-Price Competition," *Journal of Political Economy*, 76.1, 149-54.

Sutton, C. J. (1974), "Advertising, Concentration and Competition," *Economic Journal*, 84, 56-69.

Symeonidis, G. (2000), "Price Competition and Market Structure: The Impact of Cartel Policy on Concentration in the U.K.," *Journal of Industrial Economics*, 48.1, 1-26.

Varian, H. (1980), "A Model of Sales," *American Economic Review*, 70.4, 651-659.

10 Supplementary Materials

10.1 Introduction

These supplementary materials contain three parts. The first part defines a complete-information game, characterizes the associated symmetric mixed-strategy equilibrium, and shows that the distribution of advertising in this equilibrium is approximately the same as that which is induced by the pure-strategy advertising equilibrium of the incomplete-information game when production costs vary sufficiently little with respect to types. The second part considers the repeated game and extends our analysis so as to provide a more general characterization of optimal collusion. The third part considers the robustness of the results of the repeated game to a relaxation under which price selections are publicly observed by all firms.

10.2 Equilibrium in Complete-Information Game

Suppose that N firms sell a homogeneous good at a constant cost $c > 0$. A pure strategy for firm i is $A_i \in [0, r(p(c), c)]$ and A_{-i} denotes the $(N-1)$-tuple of advertising selected by other firms. The profit for firm i is

$$
\Pi_i(A_i, A_{-i}) = \begin{cases} r(p(c), c)\frac{U}{N} - A_i & \text{if } A_i < \max_{j \neq i} A_j \\ r(p(c), c)\left[\frac{U}{N} + \frac{I}{k}\right] - A_i & \begin{array}{l} \text{if } A_i \geq \max_{j \neq i} A_j \text{ and} \\ \|\{j \mid A_j = A_i\}\| = k - 1. \end{array} \end{cases}
$$

The term $r(p(c), c)$ represents $[p(c) - c]D(p(c))$. A mixed strategy for firm i is a distribution function Φ over $[\underline{A}(\Phi), \overline{A}(\Phi)]$. The profit for firm i is

$$
E_i(\Phi_i, \Phi_{-i}) = \int_{\underline{A}}^{\overline{A}} \cdots \int_{\underline{A}}^{\overline{A}} \Pi(A_i, A_{-i}) d\Phi_1 \cdots d\Phi_N,
$$

where $\overline{A}$ and $\underline{A}$ are defined below. This complete-information game has a unique symmetric mixed-strategy equilibrium, $\Phi = \Phi_i$ for all i, which is characterized as follows:

Lemma A1. *(i) There is no pure-strategy Nash equilibrium. (ii) There is a unique symmetric mixed-strategy equilibrium:*

$$
\Phi(A) = \left(\frac{A}{r(p(c), c)I}\right)^{\frac{1}{N-1}} \quad \text{with } \underline{A}(\Phi) = 0 \text{ and } \overline{A}(\Phi) = r(p(c), c)I. \tag{A1}
$$

Proof. To prove (i), assume that there are k firms that select the highest advertising A. First, suppose that $2 \leq k \leq N$. If $A < r(p(c), c)I$, then a firm can gain by raising A slightly by ε and winning all the informed consumers:

$$
r(p(c), c)\left[\frac{U}{N} + I\right] - A - \varepsilon > r(p(c), c)\left[\frac{U}{N} + \frac{I}{k}\right] - A.
$$

If $A = r(p(c), c)I$, then a firm can increase its profit by reducing A to zero and winning only the

1

uninformed consumers:

$$r(p(c),c)\frac{U}{N} > r(p(c),c)\left[\frac{U}{N} + \frac{I}{k}\right] - A.$$

Second, suppose that $k = 1$. The highest-advertising firm can raise its profit by setting $A - \varepsilon$ which is slightly above the second-highest advertising.

To prove (ii), we begin by showing that any symmetric Nash equilibrium, Φ, must satisfy (A1). To this end, we establish four findings. First, there is no mass point in Φ. If A is a mass point of Φ, then there is a positive probability of tie at A. A firm can increase its profit, if it preserves the hypothesized equilibrium strategy, except that it replaces the selection of A with the selection of $A + \varepsilon$ for small ε. Second, $\underline{A}(\Phi) = 0$. Suppose that $\underline{A}(\Phi) > 0$. If a firm chooses $\underline{A}(\Phi)$, then it wins only the uninformed consumers with probability one, since ties occur with zero probability (because of no mass point). The firm can increase its profit when it replaces the selection of $[\underline{A}(\Phi), \underline{A}(\Phi) + \varepsilon]$ with the selection of zero advertising. Third, $\overline{A}(\Phi) = r(p(c),c)I$. This result is immediate, since the profit at the top is equal to the profit at the bottom in the mixed-strategy equilibrium:

$$r(p(c),c)\left[\frac{U}{N} + I\right] - \overline{A}(\phi) = r(p(c),c)\frac{U}{N}.$$

Fourth, Φ is strictly increasing over $(\underline{A}(\Phi), \overline{A}(\Phi))$. Suppose that there is a gap (A_1, A_2) such that $\underline{A}(\Phi) < A_1 < A_2 < \overline{A}(\Phi)$ and $\Phi(A_1) = \Phi(A_2)$. Advertisements in the interval (A_1, A_2) are then selected with zero probability. For ε small, a firm would gain by replacing the selection of advertising levels in the interval $[A_2, A_2 + \varepsilon]$ with the selection of $A_1 + \varepsilon$. This deviation has the same probability of winning but uses a lower level of advertising. Given these four findings, we may conclude that, in any symmetric Nash equilibrium, Φ, and for all $A \in [0, r(p(c),c)I]$,

$$r(p(c),c)\left[\frac{U}{N} + [\Phi(A)]^{N-1}I\right] - A = r(p(c),c)\frac{U}{N}. \tag{A2}$$

This equation yields (A1). Thus, (A1) is necessarily satisfied in a symmetric Nash equilibrium. Observe next that (A1) identifies a well-defined and unique distribution function $\Phi(A)$. Lastly, we verify that Φ is a Nash equilibrium. A firm earns the same expected profit for any $A \in [\underline{A}(\Phi), \overline{A}(\Phi)]$ when all other $N-1$ firms adopt $\Phi(A)$. It cannot increase the profit by altering the distribution over the interval. Any advertising above $\overline{A}(\Phi)$ earns a lower expected profit than does $\overline{A}(\Phi)$, because $\overline{A}(\Phi)$ wins the informed consumers with probability one. Any advertising below $\underline{A}(\Phi)$ is infeasible. ∎

Purification. We consider an incomplete-information game, where production costs rise in types θ. We argue that if each firm of type θ chooses $A(\theta)$, which is the unique advertising equilibrium in the incomplete-information game, then the probability distribution induced by A is approximately the distribution of advertising in the mixed-strategy equilibrium, when the payoff relevance of types θ gets small. In the incomplete-information game, the firm of type $\theta \in [\underline{\theta}, \overline{\theta}]$ privately observes its type and has the cost $c(\theta)$. Assume that function c is differentiable and strictly increasing in θ, with $0 < c(\underline{\theta}) < c(\overline{\theta}) < p_R$, where p_R is given by $D(p_R) = 0$. The static game is the same as in the text. Then, arguing as in the proof of Proposition 1, there is a unique advertising equilibrium A which satisfies:

$$A'(\theta) = -r(p(\theta),\theta)(N-1)[1 - F(\theta)]^{N-2}f(\theta)I < 0 \text{ and } A(\overline{\theta}) = 0, \tag{A3}$$

where $r(p(\theta), \theta) = [p(c(\theta)) - c(\theta)]D(p(c(\theta)))$.

Lemma A2. *Given a constant $c \in (0, p_R)$, for any $\varepsilon > 0$, there exists $\delta > 0$ such that if $|c(\theta) - c| < \delta$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$, then the probability distribution of advertising induced by the advertising equilibrium in the incomplete-information game is ε-close to Φ_c, where Φ_c is the distribution of advertising in the mixed-strategy equilibrium of the complete-information game with constant cost c.*

   Proof. The distribution induced by $A(\theta)$ is

$$prob\left(\theta \mid A(\theta) \leq x\right) = prob\left(\theta \geq A^{-1}(x)\right) = 1 - F\left(A^{-1}(x)\right).$$

Let Φ_c denote the symmetric mixed-strategy equilibrium with costs c. Define the function A_c by

$$A_c(\theta) = \Phi_c^{-1}(\phi(\theta)), \text{ where } \phi(\theta) \equiv 1 - F(\theta).$$

Given that $\phi(\theta)$ is strictly decreasing in θ and Φ_c is strictly increasing, $A_c(\theta)$ is strictly decreasing in θ. The proof is established as a consequence of the following results. First, if each firm of type θ chooses $A_c(\theta)$, then the distribution of advertising becomes Φ_c. In other words, $A_c(\theta)$ induces the same distribution of advertising as Φ_c:

$$\begin{aligned}
prob\left(A_c(\theta) \leq x\right) &= prob\left(\Phi_c^{-1}(\phi(\theta)) \leq x\right) \\
&= prob\left(\phi(\theta) \leq \Phi_c(x)\right) \\
&= prob\left(F(\theta) \geq 1 - \Phi_c(x)\right) \\
&= prob\left(\theta \geq F^{-1}(1 - \Phi_c(x))\right) \\
&= 1 - F\left(F^{-1}(1 - \Phi_c(x))\right) \\
&= \Phi_c(x).
\end{aligned}$$

Second, $A_c(\theta)$ solves (A3) when $c(\theta) = c$. By the definition of $A_c(\theta)$, we have that

$$A_c'(\theta) = -f(\theta)/\Phi_c'(A_c(\theta)).$$

To find $\Phi_c'(A_c(\theta))$, we recall the mixed strategy (A2) and differentiate it with respect to A:

$$1 = (N-1)r(p(c), c)\left[\Phi(A)\right]^{N-2}\Phi'(A)I.$$

Replacing Φ with Φ_c, we obtain

$$\Phi_c'(A) = \frac{1}{(N-1)r(p(c), c)\left[\Phi_c(A)\right]^{N-2}I}.$$

Substituting, we thus find that

$$A_c'(\theta) = -(N-1)r(p(c), c)\left[\Phi_c(A_c(\theta))\right]^{N-2}f(\theta)I.$$

Note also that $A_c(\overline{\theta}) = \Phi_c^{-1}(1 - F(\overline{\theta})) = \Phi^{-1}(0) = 0$. Hence, when $c(\theta) = c$, $A_c(\theta)$ solves (A3). Third, if $|c(\theta) - c|$ is small, then $A(\theta)$ induces approximately the same distribution of advertising

3

as does Φ_c. This result is based on the first and second result. The function $A_c(\theta)$ induces Φ_c by the first result, and $A_c(\theta)$ approximates $A(\theta)$ when $c(\theta)$ approaches c by the second result: for any $\varepsilon > 0$, there exists $\delta > 0$ such that if $|c(\theta) - c| < \delta$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$, then $|A(\theta) - A_c(\theta)| < \varepsilon$. As $c(\theta)$ becomes closer to a constant c, the type θ becomes less payoff-relevant. Hence, the distribution of advertising induced by $A(\theta)$, $\text{prob}(\theta \mid A(\theta) \le x)$, approximates Φ_c when the payoff relevance of types θ gets small. ∎

10.3   Optimal SPPE in the Repeated Game

We now extend the analysis in the main paper that leads to Lemma 3. To this end, suppose that the entire advertising schedule is decomposed into K intervals, $[[\theta_1, \theta_2], (\theta_2, \theta_3], ..., (\theta_K, \theta_{K+1}]]$, where $\theta_1 = \underline{\theta}$ and $\theta_{K+1} = \overline{\theta}$, and $\theta_k < \theta_{k+1}$. Using Lemma 3, if the schedule A solves the No-Wars Program, then the expected profit is

$$E_\theta\left[R(\theta, \theta; A) - A(\theta)\right] = r(p(\overline{\theta}), \overline{\theta})M(\theta_K, \theta_{K+1}; A) \tag{A4}$$
$$+ \sum_{k=1}^{K} \int_{\theta_k}^{\theta_{k+1}} D(p(x))\frac{F}{f}(x)M(\theta_k, \theta_{k+1}; A)f(x)dx.$$

The interval at the top is pooling so that

$$M(\theta_K, \theta_{K+1}; A) = \frac{U}{N} + [1 - F(\theta_K)]^{N-1}\frac{I}{N}.$$

The market-share allocation function on a pooling interval is

$$M(\theta_k, \theta_{k+1}; A) = \frac{U}{N} + \sum_{j=0}^{N-1} \binom{N-1}{j}\frac{1}{j+1}[F(\theta_{k+1}) - F(\theta_k)]^j[1 - F(\theta_{k+1})]^{N-j-1}I. \tag{A5}$$

On a sorting interval, the market share allocated to type $\theta \in (\theta_k, \theta_{k+1}]$ is

$$M(\theta_k, \theta_{k+1}; A) = \frac{U}{N} + [1 - F(\theta)]^{N-1}I. \tag{A6}$$

Note that the expected market share over the entire interval is $\frac{1}{N}$:

$$\sum_{k=1}^{K} \int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A)f(\theta)d\theta = \frac{1}{N}.$$

An advertising schedule has a discontinuity (a jump) between two flat steps (pooling intervals) and between sorting and pooling intervals. The level of jump at a point is determined by the binding incentive constraint at the point.

We next ask whether an optimal SPPE adopts flat steps other than at the top. To provide a sufficient condition for an interval to take a flat step, we first show that either pooling or sorting has the same expected market share on a given interval. The market-share allocation function on a pooling (sorting) interval is denoted by $M(\theta_k, \theta_{k+1}; A^p)$ $(M(\theta_k, \theta_{k+1}; A^s))$. These functions are defined in (A5) and (A6), respectively.

Lemma A3. *Fix a partition of types* $[[\theta_1, \theta_2], (\theta_2, \theta_3], ..., (\theta_K, \theta_{K+1}]]$. *On any interval* $(\theta_k, \theta_{k+1}]$, *separating and pooling market-share allocation functions generate the same expected market share:*

$$\int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^s) f(x) dx = \int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^p) f(x) dx.$$

Proof. We follow two steps to prove that either sorting or pooling has the same expected market share:

$$\forall \theta_{k+1} \geq \theta_k, \quad \int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^p) f(x) dx = \int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^s) f(x) dx. \tag{A7}$$

First, we can immediately show that if $\theta_{k+1} = \theta_k$, then the equation holds. Second, we can show that for any $\theta_{k+1} > \theta_k$,

$$\frac{\partial \int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^p) f(x) dx}{\partial \theta_{k+1}} = \frac{\partial \int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^s) f(x) dx}{\partial \theta_{k+1}}. \tag{A8}$$

In other words, the expected market-share allocations are the same in both schemes at $\theta_{k+1} = \theta_k$, and then they increase at the same rate as θ_{k+1} rises above θ_k. The LHS of (A8) is given by

$$\frac{\partial \int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^p) f(x) dx}{\partial \theta_{k+1}} = \left[\frac{U}{N} + [1 - F(\theta_{k+1})]^{N-1} I \right] f(\theta_{k+1})$$

$$= \frac{\partial \int_{\theta_k}^{\theta_{k+1}} \left[\frac{U}{N} + [1 - F(x)]^{N-1} I \right] f(x) dx}{\partial \theta_{k+1}}$$

$$= \frac{\partial \int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^s) f(x) dx}{\partial \theta_{k+1}}.$$

The last term is the RHS of (A8). The first equality is established by tedious works of induction for $N \geq 2$. Because of the first equality, we can derive

$$\int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^p) f(x) dx = \left[[1 - F(\theta_k)]^N - [1 - F(\theta_{k+1})]^N \right] \frac{I}{N} + [F(\theta_{k+1}) - F(\theta_k)] \frac{U}{N}. \tag{A9}$$

Lemma A3 is now established. ∎

We next argue that if δ is sufficiently high, then an optimal SPPE involves a *single* flat step on any general interval over which $D(p(\theta)) \frac{F}{f}(\theta)$ is nondecreasing. We thus argue that, if a scheme on such an interval involves sorting throughout the interval, multiple flat steps, or combinations of sorting and pooling, then it is not optimal.

To build toward this result, we fix an interval $(\theta_k, \theta_{k+1}]$ and define a distribution function under a pooling scheme A^p:

$$G(\theta_k, \theta_{k+1}; A^p) \equiv \frac{\int_{\theta_k}^{\theta} M(\theta_k, \theta_{k+1}; A^p) f(x) dx}{\int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^p) f(x) dx}. \tag{A10}$$

A distribution $G(\theta_k, \theta_{k+1}; A^s)$ is analogously defined under a sorting scheme A^s. In the two distributions, the denominators are the same by Lemma A3. It then follows that $G(\theta_k, \theta_{k+1}; A^p)$ first-order stochastically dominates $G(\theta_k, \theta_{k+1}; A^s)$. Thus, for nondecreasing $D(p(\theta))\frac{F}{f}(\theta)$,

$$\int_{\theta_k}^{\theta_{k+1}} D(p(\theta))\frac{F}{f}(\theta)dG(\theta_k, \theta_{k+1}; A^p) \geq \int_{\theta_k}^{\theta_{k+1}} D(p(\theta))\frac{F}{f}(\theta)dG(\theta_k, \theta_{k+1}; A^s). \tag{A11}$$

The inequality can be rewritten to show that a flat step is preferred to a sorting scheme on an interval $(\theta_k, \theta_{k+1}]$ where $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing:

$$\int_{\theta_k}^{\theta_{k+1}} D(p(x))\frac{F}{f}(x)M(\theta_k, \theta_{k+1}; A^p)f(x)dx \geq \int_{\theta_k}^{\theta_{k+1}} D(p(x))\frac{F}{f}(x)M(\theta_k, \theta_{k+1}; A^s)f(x)dx. \tag{A12}$$

The analysis just presented establishes the first-order stochastic dominance of A^p over A^s. The argument presented uses Lemma A3, which ensures that the denominators of the distributions $G(\theta_k, \theta_{k+1}; A^p)$ and $G(\theta_k, \theta_{k+1}; A^s)$ are the same. We pause now to show that the result can be extended to the comparison between a single-step function A^p and a two-step function A^{2step} over an interval $(\theta_k, \theta_{k+1}]$. To this end, suppose that there is a jump at $\theta^* \in (\theta_k, \theta_{k+1})$, and define a distribution:

$$G(\theta_k, \theta_{k+1}; A^{2step}) \equiv \frac{\int_{\theta_k}^{\theta^*} M(\theta_k, \theta^*; A^{2step})f(x)dx + \int_{\theta^*}^{\theta} M(\theta^*, \theta_{k+1}; A^{2step})f(x)dx}{\int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^{2step})f(x)dx}. \tag{A13}$$

Observe that if the denominators of $G(\theta_k, \theta_{k+1}; A^p)$ and $G(\theta_k, \theta_{k+1}; A^{2step})$ are the same, then $G(\theta_k, \theta_{k+1}; A^p)$ first-order stochastically dominates $G(\theta_k, \theta_{k+1}; A^{2step})$. It thus suffices to show that

$$\int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^{2step})f(x)dx = \int_{\theta_k}^{\theta_{k+1}} M(\theta_k, \theta_{k+1}; A^p)f(x)dx. \tag{A14}$$

Because of (A9), the RHS of (A14) becomes

$$\left[[1 - F(\theta_k)]^N - [1 - F(\theta_{k+1})]^N\right]\frac{I}{N} + [F(\theta_{k+1}) - F(\theta_k)]\frac{U}{N},$$

and the LHS becomes

$$\int_{\theta_k}^{\theta^*} M(\theta_k, \theta^*; A^{2step})f(x)dx + \int_{\theta^*}^{\theta_{k+1}} M(\theta^*, \theta_{k+1}; A^{2step})f(x)dx$$

$$= \left[[1 - F(\theta_k)]^N - [1 - F(\theta^*)]^N\right]\frac{I}{N} + [F(\theta^*) - F(\theta_k)]\frac{U}{N}$$

$$+ \left[[1 - F(\theta^*)]^N - [1 - F(\theta_{k+1})]^N\right]\frac{I}{N} + [F(\theta_{k+1}) - F(\theta^*)]\frac{U}{N}.$$

A simplification confirms (A14). The result can be extended to any form of multiple-step functions.

The inequality (A12) indicates that there is a force in favor of pooling on an interval where $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing. It is premature, however, to conclude that optimal SPPE always entails a single pooling step on such an interval. In particular, our discussion so far has ignored the possibility that pooling on a given interval may have a negative externality on the profits for

other types on other intervals. Fortunately, however, in our model, a pooling step on one interval is not harmful to types on other intervals. The reason is that for any candidate scheme for optimal SPPE, there is an alternative scheme that has a pooling step on an interval and maintains the initial market-share allocations on other intervals. To see this, suppose that an optimal scheme, A, does not entail a single pooling step over an interval $(\theta_i, \theta_{i+1}]$ on which $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing. Our attention can be restricted to the candidate scheme that has no wars and is pooling with zero advertising at the top. We now construct an alternative scheme $\widetilde{A}$ from the top: $\widetilde{A}$ preserves the original scheme A for $\theta > \theta_{i+1}$ and $\widetilde{A}$ is pooling over $(\theta_i, \theta_{i+1}]$ and makes a parallel shift from A for $\theta \leq \theta_i$.[40] There are jumps at θ_i and θ_{i+1}. The level of jump is made such that (On-IC) is binding at each point. Observe that the alternative scheme preserves the market-share allocations of the original scheme except for the types on $(\theta_i, \theta_{i+1}]$, and that the pooling on $(\theta_i, \theta_{i+1}]$ affects (On-IC) for all the types below θ_{i+1} and their interim-stage profits. Having the same profit at the top, their interim profits take different forms of information rents. Assuming that $\widetilde{A}$ consists of K intervals, the expected profit becomes

$$
\begin{aligned}
E_\theta\left[R(\theta, \theta; \widetilde{A}) - \widetilde{A}(\theta)\right] \;=\; & r(p(\overline{\theta}), \overline{\theta})M(\theta_K, \theta_{K+1}; \widetilde{A}) \qquad\qquad\qquad\qquad\qquad\text{(A15)} \\
& + \sum_{k\neq i, k=1}^{K} \int_{\theta_k}^{\theta_{k+1}} D(p(x))\frac{F}{f}(x)M(\theta_k, \theta_{k+1}; \widetilde{A})f(x)dx \\
& + \int_{\theta_i}^{\theta_{i+1}} D(p(x))\frac{F}{f}(x)M(\theta_i, \theta_{i+1}; \widetilde{A})f(x)dx.
\end{aligned}
$$

The first two terms on the RHS are the same under the original scheme, A. The last term is in favor of pooling when $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing over $(\theta_i, \theta_{i+1}]$. This result contradicts the optimality of A and shows that $D(p(\theta))\frac{F}{f}(\theta)$ being nondecreasing over an interval is a sufficient condition for optimal SPPE to take a pooling step over the interval.

By the same token, a sorting scheme is optimal on an interval where $D(p(\theta))\frac{F}{f}(\theta)$ is decreasing, ignoring its impact on types on the other intervals.[41] We find, however, that $D(p(\theta))\frac{F}{f}(\theta)$ being decreasing over an interval is only a necessary (not a sufficient) condition for optimal SPPE to entail sorting over the interval. Suppose that $D(p(\theta))\frac{F}{f}(\theta)$ decreases over $(\theta_{i-1}, \theta_i]$ and then rises over $(\theta_i, \theta_{i+1}]$ (or is followed by a pooling step at the top). If a sorting scheme is selected over $(\theta_{i-1}, \theta_i]$ and is followed by a pooling step over the next interval, then the sorting scheme may have a negative externality on types above θ_i. If $D(p(\theta))\frac{F}{f}(\theta)$ decreases slowly on a rather short interval and rises sharply on the next interval (or if $r(p(\overline{\theta}), \overline{\theta})$ is very high), then a single flat step over the two intervals may be optimal because of the importance of pooling over the second interval.[42]

We may now summarize our findings as follows:

Lemma A4. *Allow for any F and assume that δ is sufficiently high. (i) An optimal SPPE entails a single flat step on any interval where $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing, and has a pooling step at the bottom and at the top. (ii) If a sorting scheme is ever used, it is restricted to a subset of the interval on which $D(p(\theta))\frac{F}{f}(\theta)$ is decreasing.*

---

[40] The definition of $\widetilde{A}$ is detailed in the proof of Lemma A4 below.

[41] This result is directly given when we multiply both sides of (A11) by -1.

[42] If a scheme takes a separate pooling step on the interval where $D(p(\theta))\frac{F}{f}(\theta)$ is decreasing, it is not optimal, since there is an alternative scheme that has a separate sorting step on the interval and maintains the initial market-share allocations on the other intervals.

Pooling at the top is immediate from Lemma 3 in the main paper, and $D(p(\theta))\frac{F}{f}(\theta)$ strictly increases at the neighborhood of $\underline{\theta}$ when $f(\underline{\theta}) > 0$. The result ensures that an optimal SPPE for patient firms involves a single step or multiple steps. The remainder of the proof follows.

Proof. Suppose $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing over an interval $(\theta_i, \theta_{i+1}]$. Assume further that the initial scheme A is *not* constant over this interval. We now define $\widetilde{A}$ and derive the corresponding expected profit seen in the text. Recalling the definition of A_+ and M_+ from the proof of Lemma 2 in the main paper, we define an alternative scheme $\widetilde{A}$ as

$$
\widetilde{A}(\theta) = \begin{cases} A(\theta) & \text{if } \theta > \theta_{i+1} \\ \widetilde{A}^p \equiv A_+(\theta_{i+1}) + r(p(\theta_{i+1}), \theta_{i+1})\left[M(\theta_i, \theta_{i+1}; \widetilde{A}) - M_+(\theta_{i+1}; \widetilde{A})\right] & \text{if } \theta \in (\theta_i, \theta_{i+1}] \\ \widetilde{A}(\theta_i) \equiv r(p(\theta_i), \theta_i)\left[M(\theta_i; \widetilde{A}) - M(\theta_i, \theta_{i+1}; \widetilde{A})\right] + \widetilde{A}^p & \text{if } \theta = \theta_i \\ A(\theta) - \left[A(\theta_i) - \widetilde{A}(\theta_i)\right] & \text{if } \theta < \theta_i \end{cases}
$$

The alternative scheme jumps at θ_i and θ_{i+1} such that (On-IC) is binding at each point. It preserves A above θ_{i+1}, pools over $(\theta_i, \theta_{i+1}]$ and makes a parallel shift from A by $A(\theta_i) - \widetilde{A}(\theta_i)$ below θ_i. We assume that $\widetilde{A}$ consists of K intervals. The initial scheme A may not have K intervals; it may have less than K intervals. (On-IC) for the types below θ_{i+1} and their interim profits are affected under $\widetilde{A}$. The interim profit for $\theta \in (\theta_i, \theta_{i+1}]$ is

$$
\begin{aligned}
R(\theta, \theta; \widetilde{A}) - \widetilde{A}(\theta) &= r(p(\overline{\theta}), \overline{\theta})M(\theta_K, \theta_{K+1}; \widetilde{A}) + \sum_{k=i+1}^{K} \int_{\theta_k}^{\theta_{k+1}} D(p(x))M(\theta_k, \theta_{k+1}; \widetilde{A})dx \\
&\quad + \int_{\theta}^{\theta_{i+1}} D(p(x))M(\theta_i, \theta_{i+1}; \widetilde{A})dx,
\end{aligned}
$$

where the intervals are defined to correspond to regions over which $\widetilde{A}$ entails pooling and separation, respectively. The interim profit for $\theta \in (\theta_{j-1}, \theta_j]$ $\forall j \le i$ is

$$
\begin{aligned}
R(\theta, \theta; \widetilde{A}) - \widetilde{A}(\theta) &= r(p(\overline{\theta}), \overline{\theta})M(\theta_K, \theta_{K+1}; \widetilde{A}) + \sum_{k\ne i, k=j}^{K} \int_{\theta_k}^{\theta_{k+1}} D(p(x))M(\theta_k, \theta_{k+1}; \widetilde{A})dx \\
&\quad + \int_{\theta}^{\theta_j} D(p(x))M(\theta_{j-1}, \theta_j; \widetilde{A})dx + \int_{\theta_i}^{\theta_{i+1}} D(p(x))M(\theta_i, \theta_{i+1}; \widetilde{A})dx.
\end{aligned}
$$

Since $\widetilde{A}$ preserves the initial market-share allocations except for $\theta \in (\theta_i, \theta_{i+1}]$, what is new under $\widetilde{A}$ is the last term in each interim profit. Integrating by parts, we can yield the expected profit as stated in (A15). ∎

Despite various possibilities of sorting and pooling intervals, an optimal SPPE may take a simple form. The use of sorting schemes is quite limited, and the number of pooling steps is also limited, since a separate pooling step is never used over the interval where $D(p(\theta))\frac{F}{f}(\theta)$ decreases. The simplicity of an optimal SPPE is evident when $D(p(\theta))\frac{F}{f}(\theta)$ is quasiconcave, which is likely when F is log-concave and which includes as a special case the possibility emphasized above in which $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing. If $D(p(\theta))\frac{F}{f}(\theta)$ is quasiconcave, then an optimal SPPE will be characterized by either a single pooling step with zero advertising or two pooling steps that may

or may not include a sorting interval in the middle. Consider a linear demand function $D(p) = 1 - p$ when $N = 5$, and suppose that θ is uniformly distributed over $[0, \bar{\theta}]$. Then $D(p(\theta)) \frac{F}{f}(\theta) = \frac{(1-\theta)\theta}{2}$ is concave with the maximum at 0.5. When $\bar{\theta}$ decreases from 1, pooling becomes more desirable, since profit at the top rises and the interval on which $D(p(\theta)) \frac{F}{f}(\theta)$ decreases becomes shortened. When $\bar{\theta} = 0.99$, an optimal SPPE has two flat steps that include a sorting interval approximately between 0.752 and 0.962. When $\bar{\theta} = 0.77$, an optimal SPPE has only two flat steps with a jump at 0.75. When $\bar{\theta} = 0.70$, it is a single step with zero advertising.

As we show above, if firms are sufficiently patient and $D(p(\theta)) \frac{F}{f}(\theta)$ fails to be nondecreasing over $[\underline{\theta}, \bar{\theta}]$, then the optimal SPPE is again stationary but may involve multiple pooling steps or perhaps even sorting intervals. For such cases, we can characterize the critical discount factor in the same general fashion as described in the main paper. Along a pooling step, the type that is most tempted to cheat is the lowest type on the step. If a sorting interval exists, then advertising must rise discontinuously as the type is lowered below the lowest type on the sorting interval. The lowest type on the sorting interval already advertises strictly more than all other types on this interval, and thus does not gain from a slight off-schedule increase in its advertising level. Building on this reasoning, it can be shown that the off-schedule constraint is sure to hold if the lowest type on any pooling interval does not gain from a slight increase in its advertising.

10.4 Public Price Histories

In our repeated-game analysis, we assume that each firm observes the realization of rival firms' past advertising choices but not the realization of rival firms' past pricing choices. This assumption may be appropriate in markets with complex and customer-specific pricing schemes, or when search costs are high. It also enables us to set prices at monopoly levels, so that we may use results from the static model and focus on the incentive constraints that are associated with collusion in advertising. The assumption is not always plausible, however, and we now discuss the robustness of our analysis when this assumption is relaxed.

When our repeated game is extended to allow for public price histories, each firm observes the realizations of rival firms' past advertising *and* price choices. Thus, in the extended model, a firm with cost type θ can undertake an on-schedule deviation only if it mimics the advertising *and* price selection of a firm with cost type $\widehat{\theta}$. The gain from mimicry is thus reduced, and so new equilibria exist. At the same time, our featured SPPE - in which firms pool at zero advertising and set their monopoly prices - continues to exist when price histories are public. In this equilibrium, firms simply condition their future play on the public history of advertising, and firms again set their prices at monopoly levels.

Formally, in the repeated game with public price histories, we denote a candidate advertising and pricing schedule as $(A, \widetilde{p})$, where $\widetilde{p}(\theta)$ may differ from $p(\theta)$. If a firm of cost type θ mimics the advertising and price selection of a firm of cost type $\widehat{\theta}$, then it must select $A(\widehat{\theta})$ and $\widetilde{p}(\widehat{\theta})$. To use the Relaxed Program, we let $W(\widehat{\theta}) \equiv \delta[\sup V - \bar{v}(A(\widehat{\theta}), \widetilde{p}(\widehat{\theta}); A, \widetilde{p})]$ and write the interim-stage profit as

$$\Pi(\widehat{\theta}, \theta; A, \widetilde{p}) \equiv r(\widetilde{p}(\widehat{\theta}), \theta) M(\widehat{\theta}; A) - A(\widehat{\theta}) - W(\widehat{\theta}).$$

For simplicity, assume that A and $\widetilde{p}$ are continuously differentiable except at a finite number of points where the functions may jump.

The scheme $(A, \widetilde{p}, W)$ satisfies on-schedule incentive compatibility only if two conditions hold. First, a local optimality condition must hold. Under an appropriate envelope theorem (Milgrom

and Segal, 2002), we may use $\Pi_2(\widehat{\theta}, \theta; A, \widetilde{p}) = -D(\widetilde{p}(\widehat{\theta}))M(\widehat{\theta}; A)$ to get

$$\Pi(\theta, \theta; A, \widetilde{p}) = r(\widetilde{p}(\overline{\theta}), \overline{\theta})M(\overline{\theta}; A) - A(\overline{\theta}) - W(\overline{\theta}) + \int_\theta^{\overline{\theta}} D(\widetilde{p}(x))M(x; A)dx.$$

Second, a monotonicity condition must hold. On-schedule incentive compatibility implies

$$\begin{aligned} r(\widetilde{p}(\theta), \theta)M(\theta; A) - A(\theta) - W(\theta) &\geq r(\widetilde{p}(\widehat{\theta}), \theta)M(\widehat{\theta}; A) - A(\widehat{\theta}) - W(\widehat{\theta}) \\ r(\widetilde{p}(\widehat{\theta}), \widehat{\theta})M(\widehat{\theta}; A) - A(\widehat{\theta}) - W(\widehat{\theta}) &\geq r(\widetilde{p}(\theta), \widehat{\theta})M(\theta; A) - A(\theta) - W(\theta). \end{aligned}$$

Adding the two inequalities, we find that $D(\widetilde{p}(\theta))M(\theta; A)$ must be nonincreasing in θ. As in Lemma 1, these two necessary conditions are also sufficient for $(A, \widetilde{p}, W)$ to satisfy on-schedule incentive compatibility.

We now further restrict attention to those incentive-compatible schemes $(A, \widetilde{p}, W)$ for which informed consumers are rational in using the advertising search rule.[43] With this restriction, we find that $A(\theta)$ must be nonincreasing.[44] Since informed consumers use the advertising search rule, $M(\theta; A)$ must be nonincreasing as well. Given the restriction that informed consumers are rational to use the advertising search rule, $\widetilde{p}(\theta)$ must be nondecreasing; equivalently, $D(\widetilde{p}(\theta))$ must be nonincreasing. Thus, a scheme $(A, \widetilde{p}, W)$ satisfies on-schedule incentive compatibility and is also consistent with the rational use of the advertising search rule only if $M(\theta; A)$ and $D(\widetilde{p}(\theta))$ are each nonincreasing.

Consider now the potential use of wars. When prices are public, we cannot immediately use the arguments in Lemma 2 to establish that wars are unnecessary. The reason is that incentive compatibility no longer ensures that $A(\theta) + W(\theta)$ is nonincreasing; hence, we cannot be sure that an alternative scheme defined by $\widetilde{A}(\theta) \equiv A(\theta) + W(\theta)$ would exhibit the necessary nonincreasing property.[45] In the limiting case where demand is inelastic, however, we can establish that wars are unnecessary. In that case, for an initial scheme that involves a war on a step, it is possible to eliminate the war and adjust price and advertising on that step, while ensuring that the induced advertising schedule is nonincreasing and that market shares and profits are maintained for all types.[46] When demand is elastic, however, such step-by-step maneuvers are not possible. The

---

[43] This restriction holds automatically when prices are not public, since $p(\theta)$ is strictly increasing.

[44] Assume to the contrary that $\theta > \widehat{\theta}$ and $A(\theta) > A(\widehat{\theta})$. This implies $M(\theta; A) > M(\widehat{\theta}; A)$, given that informed consumers use the advertising search rule. Since we require as well that it is rational for informed consumers to use the advertising search rule, it must be that $\widetilde{p}(\theta) \leq \widetilde{p}(\widehat{\theta})$ and hence $D(\widetilde{p}(\theta)) \geq D(\widetilde{p}(\widehat{\theta}))$. Thus, $A(\theta) > A(\widehat{\theta})$ implies $D(\widetilde{p}(\theta))M(\theta; A) > D(\widetilde{p}(\widehat{\theta}))M(\widehat{\theta}; A)$, which contradicts the requirement that $D(\widetilde{p}(\theta))M(\theta; A)$ is nonincreasing.

[45] To see that $A(\theta) + W(\theta)$ may have increasing segments, consider a two-step scheme in which A is at a high (low) level for cost types below (at or above) a critical type, θ_c. Suppose that $\widetilde{p}(\theta) = p(\theta_c)$ for types at or above θ_c while $\widetilde{p}(\theta) = p(\underline{\theta})$ for types below θ_c. Even though market share is higher for lower types, a firm with cost type θ_c may earn greater net revenue by setting its monopoly price and accepting a lower market share. On-schedule incentive compatibility would then require that $A(\theta) + W(\theta)$ is higher for higher types.

[46] Consider a two-step scheme, where $\widehat{\theta}$ represents a type on the bottom step and θ represents a type on the top step. Let θ_c denote the critical type that separates the steps. Suppose that $A(\widehat{\theta}) > A(\theta)$ and thus $M(\widehat{\theta}; A) > M(\theta; A)$. Suppose further that $A + W$ is increasing: $A(\widehat{\theta}) + W(\widehat{\theta}) < A(\theta) + W(\theta)$. Incentive compatibility is satisfied if type θ_c is indifferent between the two steps. Given that the higher step entails a lower value for M and a higher value for $A + W$, this is possible only if the higher step entails a higher price: $\widetilde{p}(\theta) > \widetilde{p}(\widehat{\theta})$. We now create a new scheme, in which $W(\theta)$ is lowered to a new value, $W_N(\theta)$, at which $A(\widehat{\theta}) + W(\widehat{\theta}) = A(\theta) + W_N(\theta) + \epsilon$, for $\epsilon > 0$ small. To maintain incentive compatibility, we adjust $\widetilde{p}(\theta)$ downward until type θ_c is again indifferent. The resulting new price $\widetilde{p}_N(\theta)$ satisfies $\widetilde{p}_N(\theta) > \widetilde{p}(\theta)$. This maneuver maintains profit for all types. We next eliminate wars and define $\widetilde{A}$ in terms of the new scheme: $\widetilde{A}(\widehat{\theta}) = A(\widehat{\theta}) + W(\widehat{\theta})$ and $\widetilde{A}(\theta) = A(\theta) + W_N(\theta)$. Note that $\widetilde{A}$ decreases with θ in the same

appeal of a price change then varies with cost type; thus, a price change on one step requires that the scheme be modified on other steps, in order for the initial market share allocation to remain incentive compatible.[47]

To allow for general demand functions, we now impose $W(\theta) \equiv 0$ and focus on stationary SPPE. Utilizing the two conditions for on-schedule incentive compatibility, we thus now analyze the No-Wars Program, in which $(A, \widetilde{p})$ is selected to maximize

$$E_\theta \left[\Pi(\theta, \theta; A, \widetilde{p})\right] = r(\widetilde{p}(\overline{\theta}), \overline{\theta})M(\overline{\theta}; A) - A(\overline{\theta}) + E_\theta \left[D(\widetilde{p}(\theta))\frac{F}{f}(\theta)M(\theta; A)\right]$$

subject to: $D(\widetilde{p}(\theta))$ and $M(\theta; A)$ are nonincreasing in θ.

A first point is that, if demand is sufficiently inelastic and F is log-concave, then zero advertising and monopoly pricing, $(A \equiv 0, p(\theta))$, solves this program. Given the requirement that $M(\theta; A)$ is nonincreasing, we may argue as before to find that profit at the top and expected information rents are maximized when $A \equiv 0$ and each type receives market share $\frac{1}{N}$. The best choice of $\widetilde{p}(\theta)$ is then the monopoly pricing function, $p(\theta)$. We conclude that our results are robust to the possibility of public price histories, if demand is sufficiently inelastic.

A second point is that pooling at zero advertising is optimal within the class of stationary SPPE in which the pricing function $\widetilde{p}(\theta)$ satisfies the further constraint that $D(\widetilde{p}(\theta))\frac{F}{f}(\theta)$ is nondecreasing. For example, if F is log-concave and all types of firms set a constant price, $\overline{p} \equiv \widetilde{p}(\theta) \geq \overline{\theta}$, then the optimal advertising schedule entails pooling at zero advertising. Firms may set a constant price for a variety of (unmodeled) reasons, including resale price maintenance requirements and customer market concerns. In fact, when these reasons apply and a constant price is used, we can argue as in Lemma 2 and show that wars are not useful (i.e., the restriction to stationary SPPE is without loss of generality, when price is constant). Thus, in a modified game where firms must use a constant price, if F is log-concave, the optimal SPPE for patient firms entails pooling at zero advertising. Of course, in the case where price is exogenously fixed at $\overline{p}$, it is immaterial whether or not price is public.

Our third point is that robust forces remain in favor of pooling in advertising, even for general demand functions.[48] A simple way to make this point is to consider any scheme $(A, \widetilde{p})$ in which $A(\theta)$ is strictly decreasing and thus entails sorting over $[\underline{\theta}, \overline{\theta}]$. We may then consider an alternative pooling scheme $(A^*, \widetilde{p}^*)$ in which $\widetilde{p}^* \equiv \overline{p}^*$ and $A^* \equiv 0$ over $[\underline{\theta}, \overline{\theta}]$. The level of price $\overline{p}^*$ is determined to satisfy

$$\int_{\underline{\theta}}^{\overline{\theta}} D(\overline{p}^*)\frac{1}{N}f(x)dx = \int_{\underline{\theta}}^{\overline{\theta}} D(\widetilde{p}(x))\left[\frac{U}{N} + [1 - F(x)]^{N-1}I\right]f(x)dx$$

We now define a distribution function under $(A^* \equiv 0, \overline{p}^*)$:

$$G(\theta; A^*, \overline{p}^*) \equiv \frac{\int_{\underline{\theta}}^{\theta} D(\overline{p}^*)\frac{1}{N}f(x)dx}{\int_{\underline{\theta}}^{\overline{\theta}} D(\overline{p}^*)\frac{1}{N}f(x)dx}.$$

A distribution $G(\theta; A, \widetilde{p})$ is analogously defined under $(A, \widetilde{p})$. Since $G(\theta; A^*, \overline{p}^*)$ first-order stochas-

---

way as did A; hence, $\widetilde{A}$ generates the same market share allocation as did A.

[47] For related reasons, Athey, Bagwell and Sanchirico (2004) are also unable to eliminate wars when demand is elastic.

[48] Our discussion here builds on Athey, Bagwell and Sanchirico (2004).

tically dominates $G\left(\theta; A, \widetilde{p}\right)$, we can show that, for nondecreasing $\frac{F}{f}(\theta)$,

$$\int_{\underline{\theta}}^{\overline{\theta}} \frac{F}{f}(x)D(\overline{p}^*)\frac{1}{N}f(x)dx \geq \int_{\underline{\theta}}^{\overline{\theta}} \frac{F}{f}(x)D(\widetilde{p}(x))\left[\frac{U}{N} + [1-F(x)]^{N-1}I\right]f(x)dx.$$

Thus, when F is log-concave, any scheme in which advertising entails sorting over the support generates lower expected information rents than does an alternative pooling scheme in which all types select zero advertising.

10.5   References

Athey, S., K. Bagwell and C. Sanchirico (2004), "Collusion and Price Rigidity," *Review of Economic Studies*, 71: 317-49.

Milgrom, P. and I. Segal, "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70.2, 583-601.