4-2010

Advertising Collusion in Retail Markets

Kyle BAGWELL
Stanford University

Gea M. LEE
Singapore Management University, gmlee@smu.edu.sg

SMU ECONOMICS & STATISTICS WORKING PAPER SERIES

Advertising Collusion in Retail Markets

Kyle Bagwell and Gea M. Lee

April 2010

Paper No. 03-2010

Advertising Collusion in Retail Markets

Kyle Bagwell and Gea M. Lee

December 21, 2009

Abstract

We analyze non-price advertising by retail firms, when the firms are privately informed about their respective costs of production. In a static advertising game, an advertising equilibrium exists in which lower-cost firms select higher advertising levels. In this equilibrium, informed consumers rationally employ an advertising search rule in which they buy from the highest-advertising firm, since lower-cost firms also select lower prices. In a repeated advertising game, colluding firms face a tradeoff: the use of advertising can promote productive efficiency but only if sufficient current or future advertising expenses are incurred. At one extreme, if firms pool at zero advertising, they sacrifice productive efficiency but also eliminate current and future advertising expenses. Focusing on symmetric perfect public equilibria for the repeated advertising game, we establish conditions under which optimal collusion entails pooling at zero advertising. More generally, full or partial pooling is observed in optimal collusion. Such collusive agreements reduce consumer welfare, since they restrict informed consumers' ability to locate the lowest available price in the market.

*Bagwell: Donald L. Lucas Professor of Economics, Stanford University; Senior Fellow at SIEPR and SCID; and NBER Research Associate. Lee: School of Economics, Singapore Management University. We thank two anonymous referees, Yeon-Koo Che, Pauline Ippolito, Jeong-Yoo Kim, Massimiliano Landi, Laurent Linnemer, Albert Ma and John Vickers for helpful comments. We also thank seminar participants at EARIE, ENPC-CERAS, EUI, the Far Eastern Meeting of the Econometric Society, Singapore Management University, and Southern Methodist University for helpful discussions. The material in this paper was originally contained in our earlier paper entitled "Advertising and Collusion in Retail Markets," dated June 2008. The dynamic analysis from the earlier paper is now found here, and the static analysis from the earlier paper is now found in our companion paper (Bagwell and Lee, 2009). Lee gratefully acknowledges Singapore Management University research fund.

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. In his classic study of the retail eyeglass industry in the U.S. in the 1960s, Benham (1972) compares transaction prices under different legal systems: prices were higher in states that prohibited all advertising than in states that had no restrictions on advertising; in addition, prices were only slightly higher in states that allowed just non-price advertising than in states that also allowed price advertising. Apparently, the ability to advertise even in only a non-price form is sometimes associated with lower prices. Similar findings are reported by Cady (1976) in his analysis of the U.S. retail market for prescription drugs in 1970. This work suggests the possibility that retail firms might sometimes gain if they are able to limit the use of non-price advertsing. Of course, retail firms directly achieve an agreement to limit advertising when a state law prohibits advertising.[3] In the absence of such a state law, retail firms may interact repeatedly and seek to achieve a self-enforcing agreement to limit advertising.

Bagwell and Ramey (1994a) offer an equilibrium interpretation of Benham's findings. They develop a complete-information model of retail competition, in which some consumers are "informed"

[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).

[3] Alternatively, retail firms might achieve such an agreement by forming a professional association that imposes advertising restrictions on its members. The FTC has argued that anti-competitive effects may be associated with price and non-price advertising restrictions imposed by a professional association. See *California Dental Association v. Federal Trade Commission* (1999).

and can identify the highest-advertising firm, while other consumers are "uninformed" and do not observe advertising levels. All consumers possess downward-sloping demand curves and must visit a firm in order to observe its price. Bagwell and Ramey focus on two kinds of equilibria. In a random equilibrium, consumers ignore advertising and choose firms at random. Consequently, firms do not advertise, and they enjoy symmetric market shares. In an advertising equilibrium, informed consumers use the advertising search rule, whereby they go to the highest-advertising firm. In response to this consumer behavior, firms employ a symmetric mixed strategy that pairs higher advertising choices with greater investments in cost reduction and thus lower prices. Informed consumers are then rational in using the advertising search rule. For a fixed number of firms, expected profit is higher in the random equilibrium, because advertising expenses are thereby avoided. In a free-entry model, their findings regarding the relationship between non-price advertising and average transaction prices are broadly consistent with the empirical patterns that Benham reports, if the random equilibrium is associated with a setting in which advertising is legally banned.

In this paper, we modify the Bagwell-Ramey model in two key respects. First, we assume that firms have private information as to their respective production costs. In particular, we consider an incomplete-information model with a continuum of possible cost types, where cost types are iid across firms. In the corresponding static game, an advertising equilibrium exists in which lower-cost firms advertise more and price lower than do higher-cost firms. Informed consumers are again rational in using the advertising search rule. The advertising equilibrium may then be compared with the random equilibrium in which no firm advertises and consumers pick a firm at random.[4] Second, we assume that firms interact repeatedly over an infinite horizon, where advertising selections are publicly observed by firms and 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 particular comparison between the random and advertising equilibria.

Assuming that informed consumers use the advertising search rule in each period, we focus on the symmetric perfect public equilibria (SPPE) of our repeated advertising game. For this class of equilibria, our goal is to characterize the optimal form of collusion in advertising among firms.[5] We note that the SPPE solution concept includes a wide range of 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

[4] In our companion paper (Bagwell and Lee, 2009), we compare the advertising equilibrium with the random equilibrium, both when the number of firms is fixed and when the number of firms is endogenous under free entry, and we thereby consider the short- and long-run implications of advertising competition for consumer surplus, firm profit and social welfare. We also analyze a benchmark model of price competition and compare the corresponding pricing equilibrium with the advertising equilibrium.

[5] 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.

the form of step functions. A further possibility is that firms implement non-stationary SPPE, in which they move between cooperative and war phases in their advertising conduct.

When firms collude in private-information settings, two kinds of incentive constraints arise.[6] 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 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 any demand function and for any distribution function of cost types. We thus confirm at a general level that future advertising wars are a redundant instrument: firms cannot achieve higher profits with a non-stationary SPPE than with a stationary SPPE. 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 an optimal SPPE for sufficiently patient firms entails pooling at zero advertising for all cost types in all periods. We also strengthen this finding and establish that, under the same conditions, *any* optimal SPPE is stationary and entails pooling at zero advertising by all cost types in all periods. Thus,

[6] The discussion here follows Athey et al. (2004) and Athey and Bagwell (2001).

while our SPPE solution concept allows for a wide range of behaviors, we show that important conditions exist under which advertising behavior in any optimal SPPE of our repeated game takes a remarkably simple form: along the equilibrium path, no firm advertises in any period.

When firms collude in this way, the welfare of consumers is reduced below that which they enjoy in the advertising equilibrium of the static game. Intuitively, in our model, the induced distribution of posted prices is independent of the advertising selections of firms. This means that uninformed consumers enjoy the same consumer surplus whether or not firms eliminate non-price advertising. In the advertising equilibrium, however, informed consumers use non-price advertising to infer the identity of the lowest-cost, and thus the lowest-price, firm in the market. The average *transaction* price is thus increased when firms collude and eliminate non-price advertising. Collusion of this kind thus acts to reduce the welfare of informed consumers.

We emphasize that the characterization of optimal collusive conduct described above 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, optimal SPPE may not entail zero advertising by all types. We thus also characterize optimal SPPE behavior under general demand functions. Requiring sufficiently patient firms, we establish three additional findings.

First, for any demand function, if the support of possible cost types is sufficiently small, then any optimal SPPE entails pooling at zero advertising by all types in all periods. Second, for any demand function and for any distribution function of cost types, any optimal SPPE involves at least partial pooling; in particular, any optimal SPPE entails pooling at the bottom and at the top (i.e., on intervals of cost types adjoining the lowest-cost and highest-cost types).[7] This second finding ensures that, under general conditions, an optimal SPPE for patient firms strictly improves upon the repeated use of the static advertising equilibrium in which advertising is strictly decreasing in cost type. Third, for a large family of demand and distribution functions, we show that any optimal SPPE uses at most two pooling intervals: it is characterized by either one pooling step at zero advertising or two pooling steps at the bottom and at the top, with or without an intermediate sorting interval. Firms limit the number of pooling steps to diminish advertising expenses. Again, such collusive agreements harm consumer welfare, since they restrict informed consumers' ability to locate the lowest available price in the market.

Our analysis of the repeated advertising game is closely related to work by Athey et al. (2004).[8] They consider a repeated game in which firms have private cost shocks and collude in pricing.

[7] When an optimal SPPE entails positive advertising for some cost types, we may generate the associated payoffs using a stationary or non-stationary SPPE. The reason is that firms may then allocate advertising expenses across periods, because of the substitutability between current-period advertising and future advertising wars.

[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, 2009).

The game considered by Athey et al. may be thought of as a repeated first-price (procurement) auction, while the repeated advertising game that we analyze here is analogous to a repeated all-pay auction. In their paper, when the distribution of cost types is log-concave, if demand is sufficiently inelastic and firms are sufficiently patient, then firms always select the same price, regardless of their respective cost types, along the equilibrium path of any optimal SPPE. As described above, we establish a similar finding in our model of collusion in advertising. As well, we report that any optimal SPPE for patient firms is stationary and entails pooling at zero advertising, even for elastic demand functions, if the support of possible cost types is sufficiently small.[9] For a large family of demand and distribution functions, we also show that an optimal SPPE entails at most two pooling intervals. Finally, Athey et al. also show that, if demand is sufficiently inelastic, then an optimal SPPE exists that is stationary. In our model of collusion in advertising, for general demand functions, an optimal SPPE exists that is stationary.

In other related work, Peters (1984) and LeBlanc (1998) consider the effects of a prohibition on price advertising in models where each firm is privately informed about its production cost. By contrast, here we emphasize that firms can achieve a self-enforcing restriction on non-price advertising. Also, Bagwell and Ramey (1994b) consider a duopoly model in which one firm has private information as to whether its costs are high or low. In a static setting, they show that non-price advertising may be used to signal low costs and thus low prices. In the current paper, by contrast, we adopt a continuum-type model in which all firms are privately informed as to their costs. In a dynamic setting where restrictions on non-price advertising must be self-enforced, we show that firms often have incentive to restrict the use of non-price advertising.

The paper is organized as follows. Section 2 contains the static advertising game. The repeated game is examined in Section 3. Optimal collusion for patient firms is characterized in Section 4. Section 5 characterizes the critical discount factor above which optimal SPPE entail pooling at zero advertising or two pooling intervals. Section 6 concludes. In the Appendix, we discuss the robustness of our analysis and provide additional proofs.

2 The Static Advertising Game

We begin with a static game in which 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 two kinds of equilibria, advertising and random equilibria, and compare the expected profits earned by firms under these two equilibria. Our analysis of the advertising game of the static model is developed further in our companion paper (Bagwell and Lee, 2009).

[9] We also consider the case of a uniform distribution of types and a demand function whose elasticity is constant and above unity. If the elasticity of demand does not exceed a critical level, then any optimal SPPE for patient firms is again stationary and entails pooling at zero advertising.

2.1 The Model

We assume $N \geq 2$ ex ante identical firms. The firms compete for sales in a homogeneous-good market, and each firm i is privately informed of its unit cost level θ_i. Firm i's cost type θ_i is drawn from the support $[\underline{\theta}, \overline{\theta}]$ according to the twice-continuously differentiable distribution function, $F(\theta)$, where $\overline{\theta} > \underline{\theta} \geq 0$. Cost types are iid across firms. We define the density as $f(\theta) \equiv F'(\theta)$, where $f(\theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$. After firms observe their individual cost types, the firms simultaneously choose their prices and levels of advertising. We follow Bagwell and Ramey (1994a) and assume that advertising is a dissipative expense that does not directly affect consumer demand.

The market contains a unit mass of consumers. Each consumer has a twice-continuously differentiable demand function $D(p)$ that satisfies $D(p) > 0 > D'(p)$ over the relevant range of prices p. We assume that prices cannot be directly communicated in the market; in particular, consumers cannot observe prices prior to picking a firm to visit and from which to purchase. Consumers are divided into two groups. A fraction I of consumers are informed in the sense that they observe firms' advertising expenses.[10] Based on this information, informed consumers form beliefs as to firms' cost types and employ a visitation (search) strategy. For instance, informed consumers may use an *advertising search rule*, in which a consumer goes to the highest-advertising firm.[11] The remaining fraction $U = 1 - I$ of consumers do not observe advertising expenditures and are uninformed. Uninformed consumers may adopt a *random search rule*, whereby a consumer randomly chooses which firm to visit.

We now define the following *advertising 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. Observe that a consumer can visit only one firm.[12] As we explain below, this assumption simplifies our analysis, since it ensures that each firm chooses the monopoly price that is associated with its cost type for any sales it makes.

For the advertising game of the static model, we are interested in Perfect Bayesian Equilibria. We impose two additional requirements on our solution concept. First, we restrict attention to equilibria in which consumers do not condition their visitation decision on firms' "names." Thus, uninformed consumers must use the random search rule, and, for any given vector of firm advertising levels, informed consumers must treat symmetrically any two firms which advertise at the same level. We note that informed consumers satisfy this requirement when they use the advertising search rule. Second, we restrict attention to equilibria in which firms use symmetric pricing and advertising strategies. Observe that the random search rule is indeed an optimal search strategy

[10] It is not essential that informed consumers observe all advertising expenditures. All of our results hold if informed consumers observe only the identity of the highest-advertising firm(s).

[11] If more than one firm advertises at the highest level, then the advertising search rule requires that informed consumers choose randomly among the highest-advertising firms.

[12] In our companion paper (Bagwell and Lee, 2009), we develop a modified advertising model in which consumers can undertake costly sequential search and firms choose advertising levels and prices. We establish the existence of an advertising equilibrium and show that the possibility of sequential search serves to strengthen our main finding that firms may achieve higher expected profit when they restrict the use of advertising.

for uninformed consumers, when firms use symmetric pricing strategies.

Using our symmetry requirement, we can define a pure advertising strategy for firm i as a function $A(\theta_i)$ that maps from the set of cost types $[\underline{\theta}, \overline{\theta}]$ to the set of possible advertising expenditures $\mathbb{R}_+ \equiv [0, \infty)$. Abusing notation somewhat, let the vector $\mathbf{A}(\boldsymbol{\theta}_{-i})$ denote the advertising selections of firms other than i when these firms all use the schedule A and their cost types are given by the $(N-1)$-tuple $\boldsymbol{\theta}_{-i}$. For any given search rule used by informed consumers, firm i's market share is determined by the vector of advertising levels selected by firm i and its rivals. Thus, the market share for firm i maps from $\mathbb{R}_+^N$ to $[0,1]$ and in equilibrium may be represented as $m(A(\theta_i), \mathbf{A}(\boldsymbol{\theta}_{-i}))$.[13] Note that, under our first requirement above, firm i's market share is not indexed by i and thus does not depend on firm i's name. Thus, if firm i has cost type θ_i, advertises at level $A(\theta_i)$ and anticipates that its rivals employ the strategy A to determine their advertising levels, then its interim-stage market share is given by $M(A(\theta_i); A) \equiv E_{\boldsymbol{\theta}_{-i}}[m(A(\theta_i), \mathbf{A}(\boldsymbol{\theta}_{-i}))]$.

We next define a firm's expected profit. Let $r(p, \theta) \equiv (p - \theta)D(p)$ denote a firm's net revenue (excluding advertising expenses) when it has cost type θ, sets the price p and sells to the entire unit mass of consumers. We assume $r(p, \theta)$ is strictly concave in p with a unique maximizer $p(\theta) = \arg\max_p r(p, \theta)$. The monopoly price $p(\theta)$ then strictly increases in θ whereas $r(p(\theta), \theta)$ strictly decreases in θ. We also assume $p(\overline{\theta}) > \overline{\theta}$, so that the price "at the top" has a positive margin. Using our requirement that all consumers, and specifically uninformed consumers, treat all firms symmetrically, we conclude that all firms must receive positive expected market share. In the equilibria upon which we focus, therefore, each firm must select the monopoly price given its cost type. We may thus embed the monopoly price into the revenue function and define the interim-stage net revenue for firm i by $R(A(\theta_i), \theta_i; A) \equiv r(p(\theta_i), \theta_i)M(A(\theta_i); A)$. We further simplify our notation by ignoring subscript i. If a firm of type θ picks an advertising level $A(\widehat{\theta})$ when its rivals employ the strategy A to determine their advertising levels, then its interim-stage profit is

$$\begin{aligned} \Pi(A(\widehat{\theta}), \theta; A) &\equiv r(p(\theta), \theta)M(A(\widehat{\theta}); A) - A(\widehat{\theta}). \\ &\equiv R(A(\widehat{\theta}), \theta; A) - A(\widehat{\theta}). \end{aligned} \tag{1}$$

With our additional requirements embedded, we now define an *equilibrium* as an advertising strategy A, a belief function and search rules for consumers that collectively satisfy three remaining conditions. First, given the market share function, m, that is induced by consumers' search rules, the advertising strategy A is such that, for all θ, $A(\theta) \in \arg\max_a [R(a, \theta; A) - a]$.[14] Second, given an observed advertising level a by a firm, informed consumers use Bayes' Rule whenever possible (i.e., whenever $a = A(\theta)$ for some $\theta \in [\underline{\theta}, \overline{\theta}]$) in forming their beliefs as to that firm's cost type θ

[13] For example, if all consumers use the random search rule, then $m(A(\theta_i), \mathbf{A}(\boldsymbol{\theta}_{-i})) = \frac{1}{N}$. If instead the uninformed consumers use the random search rule while the informed consumers use the advertising search rule, then $m(A(\theta_i), \mathbf{A}(\boldsymbol{\theta}_{-i})) = I + \frac{U}{N}$ if $A(\theta_i) > A(\theta_j)$ for all $j \neq i$, while $m(A(\theta_i), \mathbf{A}(\boldsymbol{\theta}_{-i})) = \frac{U}{N}$ if $A(\theta_i) < A(\theta_j)$ for some $j \neq i$. For this latter set of consumer search strategies, if firm i ties with $k-1$ other firms for the highest advertising level, then $m(A(\theta_i), \mathbf{A}(\boldsymbol{\theta}_{-i})) = \frac{I}{k} + \frac{U}{N}$.

[14] Notice that $A(\theta)$ must be an optimal choice for a firm with type θ in comparison to advertising deviations that are "on-schedule" (i.e., a such that $a = A(\widehat{\theta}) \neq A(\theta)$ for some $\widehat{\theta} \in [\underline{\theta}, \overline{\theta}]$) as well as "off-schedule" (i.e., a such that $a \neq A(\theta)$ for any $\theta \in [\underline{\theta}, \overline{\theta}]$).

and thus price $p(\theta)$. Third, for any observed vector of advertising levels, given their beliefs, the informed consumers' search rule directs them to the firm or firms with the lowest expected price.

For a given equilibrium, a firm of type θ advertises at level $A(\theta)$. We can thus express the firm's expected revenue and profit, respectively, as $E_\theta R(A(\theta), \theta; A)$ and $E_\theta[R(A(\theta), \theta; A) - A(\theta)]$, where the implicit market share functions are determined by the equilibrium search rules of informed consumers. In the next subsection, we restrict attention to equilibria in which consumer use particular search rules. Equilibrium market share functions may then be explicitly and simply represented.

2.2 Advertising and Random Equilibria

In this subsection, we establish the existence of two kinds of equilibria. In an *advertising equilibrium*, informed consumers use the advertising search rule. 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 first consider advertising equilibria. In such an equilibrium, firms use an advertising strategy $A(\theta)$, informed consumers use the advertising search rule, and uninformed consumers are randomly distributed across all N firms. Since $r(p(\theta), \theta)$ is strictly decreasing, lower-cost firms enjoy market share expansion more than do higher-cost firms. As the cost of advertising at any level is independent of a firm's cost type, we can thus easily show that equilibrium interim-stage market share, $M(A(\theta); A)$, must be nonincreasing in a firm's cost type, θ.[15] This implies in turn that $A(\theta)$ is also nonincreasing in θ, since at the interim stage no firm would be willing to advertise more in order to receive (weakly) less market share. Further, given the advertising search rule, $A(\theta)$ 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, which implies that $M(A(\theta); A) = \frac{U}{N} + [1 - F(\theta)]^{N-1} I$. Given $M(A(\overline{\theta}); A) = \frac{U}{N}$, a firm with type $\overline{\theta}$ will select zero advertising, and so $A(\overline{\theta}) = 0$.

These necessary conditions for an advertising equilibrium are developed in further detail in our companion paper (Bagwell and Lee, 2009). We establish there also the following existence result:

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

The advertising equilibrium acts as a fully sorting (separating) mechanism: firms truthfully reveal their cost types along the downward-sloping advertising schedule. Informed consumers rationally employ the advertising search rule, since the lowest-cost firm advertises the most and also

[15] This discussion reflects the underlying single-crossing property that holds in the model. When a firm increases its advertising level, it confronts a trade off between the larger advertising expense, a, and the consequent higher expected market share, $M(a; A)$. Holding the interim-stage profit constant, the slope $da/dM(a; A)$ is given by $r(p(\theta), \theta)$, which is strictly decreasing in θ.

offers the lowest price. Thus, ostensibly uninformative advertising directs market share to the lowest-cost supplier and promotes productive efficiency.

We next consider the random equilibrium, wherein all consumers use the random search rule and thus divide up evenly across firms. Given the random search rule, each firm receives an equal share, $\frac{1}{N}$, of the unit mass of consumers. Thus, $M(A(\theta); A) = \frac{1}{N}$ in a random equilibrium. Each firm thus chooses zero advertising in a random equilibrium, since even informed consumers are unresponsive to advertising. In addition, when firms pool and do not advertise, the random search rule is a best response for each consumer.[16] The random equilibrium thus exists and takes the form of a pooling equilibrium.

Bagwell and Lee (2009) compare expected consumer surplus in these two equilibria. Given that the induced distribution of monopoly prices is not altered across the two equilibria, uniformed consumers expect the same consumer surplus whether the advertising or random equilibrium is anticipated. Informed consumers, however, expect strictly higher consumer surplus in the advertising equilibrium than in the random equilibrium. The key point is that, in the advertising equilibrium, informed consumers can infer the identity of the lowest-cost, and thus the lowest-price, firm.

Bagwell and Lee also compare the expected profits earned by firms in these equilibria. The comparison is subtle: the advertising equilibrium achieves productive efficiency while the random equilibrium does not; however, the random equilibrium also avoids all advertising expenses. They show that firms make a strictly higher expected profit in the random equilibrium than in the advertising equilibrium, if F is log-concave and demand is sufficiently inelastic or if the support of possible cost types is sufficiently small. This result suggests that important circumstances exist under which retail firms would benefit from a restriction on non-price retail advertising. 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 such a restriction on advertising if advertising were legally prohibited. Finally, even if advertising is legal and informed consumers are responsive to advertising, firms may be able to eliminate advertising as part of an optimal self-enforcing collusive agreement. In our analysis of the repeated game below, generalizing beyond the particular comparison between the advertising and random equilibria, we confirm this possibility by showing that firms may prefer zero advertising to *any* other self-enforcing advertising scheme. Such a collusive agreement, however, reduces expected consumer surplus by eliminating the ability of informed consumers to locate the lowest available price in the market.

[16] 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 static model of advertising in which one firm has two possible cost types. In the repeated game that we analyze below, the random equilibrium is achieved as a self-enforcing ban on advertising in which a deviation from zero advertising would cause a future advertising war.

3 The Repeated Advertising Game

We consider next a repeated game in which firms select advertising levels and are privately informed with respect to their realized cost levels in each period. We assume that informed consumers use the advertising search rule. In this section, we define the repeated game and present some programs that are useful in the next section where we characterize optimal collusion for firms.

3.1 The Model

We now define the repeated game. In each of an infinite number of periods, firms play the static advertising 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 explained 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.[17]

The vectors of cost types, advertising schedules and realized advertisements at date t are denoted $\boldsymbol{\theta}_t \equiv (\theta_{it}, \boldsymbol{\theta}_{-it})$, $\mathbf{A}_t \equiv (A_{it}, \mathbf{A}_{-it})$ and $\mathbf{a}_t \equiv (a_{it}, \mathbf{a}_{-it})$. 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:

$$u_i(\{\boldsymbol{\theta}_t, \mathbf{A}_t\}_{t=1}^{\infty}) = \sum_{t=1}^{\infty} \delta^{t-1} \left[r\left(p\left(\theta_{it}\right), \theta_{it}\right) m_i(\mathbf{a}_t) - a_{it} \right], \tag{2}$$

where $a_{it} = A_{it}(\theta_{it})$ and $\delta \in (0,1)$ denotes the common discount factor for firms. 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.[18] With this assumption, the

[17] In the Appendix, we discuss the robustness of our analysis when this assumption is relaxed. We argue there that forces in favor of pooling remain, even when prices are public.

[18] In practice, production costs may consist of several components that are private. Some components, such as the price of certain raw materials or the productivity of some factors, may fluctuate in a transitory way, whereas other components, such as the details of long-term contracts with suppliers, may have a more persistent influence on

repeated game takes a recursive structure.

As our solution concept, we employ Perfect Public Equilibrium (Fudenberg et al., 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 $\{A \mid A : [\underline{\theta}, \overline{\theta}] \rightarrow \mathbb{R}_+\}$. For simplicity, we assume that any stage-game strategy A is continuously differentiable except at perhaps a finite number of points where A jumps. 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 et al. (1986, 1990), 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 Abreu et al., 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 $v : \mathbb{R}_+^N \rightarrow \mathbb{R}$. 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 $\overline{v}(A(\theta); A) \equiv E_{\boldsymbol{\theta}_{-i}}[v(A(\theta), \mathbf{A}(\boldsymbol{\theta}_{-i}))]$, where $\mathbf{A}(\boldsymbol{\theta}_{-i})$ denotes the $(N-1)$-tuple of advertising selections by firms other than i when these firms all use the schedule A. 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 $E_\theta \left[\Pi(A(\theta), \theta; A) + \delta\overline{v}(A(\theta); A)\right]$.

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

production costs. Our assumption of transitory shocks simplifies the analysis considerably, since otherwise a firm's current advertising choice could signal its cost and thereby affect the beliefs that rival firms carry into the following period. Athey and Bagwell (2008) consider a model of price collusion in the case where production costs are persistent over time and privately observed.

function v to maximize

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

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 et al. (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: $\forall\widehat{\theta}\neq\theta$,

$$\begin{aligned}&\forall\boldsymbol{\theta}_{-i},\ v(A(\widehat{\theta}),\mathbf{A}(\boldsymbol{\theta}_{-i}))\in V\\&\forall\theta,\ \Pi(A(\theta),\theta;A)+\delta\overline{v}(A(\theta);A)\geq\Pi(A(\widehat{\theta}),\theta;A)+\delta\overline{v}(A(\widehat{\theta});A)\end{aligned}$$

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

$$\begin{aligned}&\forall\boldsymbol{\theta}_{-i},\ v(\widehat{a},\mathbf{A}(\boldsymbol{\theta}_{-i}))\in V\\&\forall\theta,\ \Pi(A(\theta),\theta;A)+\delta\overline{v}(A(\theta);A)\geq\Pi(\widehat{a},\theta;A)+\delta\overline{v}(\widehat{a};A).\end{aligned}$$

Following Athey et al. (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 $v(A(\widehat{\theta}),\mathbf{A}(\boldsymbol{\theta}_{-i}))\in V$ 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. Let $\Pi(\widehat{\theta},\theta;A)\equiv\Pi(A(\widehat{\theta}),\theta;A)$, $M(\widehat{\theta};A)\equiv M(A(\widehat{\theta});A)$ and

$R(\widehat{\theta}, \theta; A) \equiv R(A(\widehat{\theta}), \theta; A)$. We 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 choice of $A(\widehat{\theta})$. 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:

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

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 2. (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 et al. (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 (On-IC) 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.[19] 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 2. 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)*

$$R(\theta,\theta;A) - A(\theta) - W(\theta) = R(\overline{\theta},\overline{\theta};A) - A(\overline{\theta}) - W(\overline{\theta}) + \int_{\theta}^{\overline{\theta}} D(p(x))M(x;A)dx. \tag{3}$$

The proof of this result is standard in the mechanism-design literature and is therefore omitted.[20] 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 stated in Proposition 1. Further, if firms strictly prefer pooling at zero advertising to using the advertising equilibrium of the static game and they 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

[19] We show below in Lemma 2 that A^* achieves strictly higher expected profit than does the advertising equilibrium of the static game.

[20] To confirm that (On-IC) implies that $M(\theta; A)$ is nonincreasing, we may fix any two types (say, θ_1 and θ_2), express the two (On-IC) constraints under which a firm with one type does not gain from mimicking the behavior assigned to the other type, and then add the two constraints. Given the consumer search rules, $M(\theta; A)$ is nonincreasing if and only if $A(\theta)$ is nonincreasing. A local optimality condition, $\Pi_1(\widehat{\theta},\theta;A) = 0$ for $\widehat{\theta} = \theta$, must also hold, and the application of an appropriate envelope theorem (Milgrom and Segal, 2002) thus yields (3). Together, the two conditions are sufficient for (On-IC), due to the single-crossing property of the model.

generally, stationary SPPE may entail advertising schedules with intervals of pooling as well as intervals of separation.

4 Optimal Collusion for Patient Firms

In this section, we characterize optimal SPPE, assuming firms are sufficiently patient so that off-schedule constraints hold. We report our findings in five steps. First, we show that equilibrium-path wars are not necessary for optimal SPPE. Second, using Proposition 2, we report conditions under which an optimal SPPE exists that is stationary, wherein firms pool at zero advertising in all periods. Third, we show that the same conditions for the second finding ensure that *any* optimal SPPE is stationary and involves pooling at zero advertising in all periods. Fourth, in a more general setting, we show that any optimal SPPE involves at least partial pooling. Fifth, we show that an optimal SPPE with partial pooling may involve quite a small number of pooling intervals.

4.1 No Wars

In this subsection, we establish a substitutability between current advertising and future wars, and we thereby conclude that the relaxed program has a no-war solution, $(A^*, W^* \equiv 0)$. Accordingly, for sufficiently patient firms, we establish that wars are not necessary for optimal SPPE. When firms are sufficiently patient, we may thus use the program specified in Proposition 2 to characterize the advertising schedule that is used in an optimal and stationary SPPE.

Suppose that a scheme (A, W) satisfies (On-IC) in the Relaxed Program. Then, we say that an alternative scheme (A^*, W^*) is *point-wise equivalent* to (A, W) if the scheme satisfies (On-IC) and preserves the market-share allocation and interim-stage profit: $\forall\theta$,

$$M(\theta; A^*) = M(\theta; A) \text{ and } R(\theta, \theta; A^*) - A^*(\theta) - W^*(\theta) = R(\theta, \theta; A) - A(\theta) - W(\theta). \qquad (4)$$

We now establish a substitutability between current advertising and future wars: for any (A, W) that satisfies (On-IC), we can set $A^*(\theta) \equiv A(\theta) + W(\theta)$ and construct a no-war scheme $(A^*, W^* \equiv 0)$ that is point-wise equivalent to (A, W).

Proposition 3. (Substitutability) *Assume that (A, W) satisfies (On-IC) in the Relaxed Program. A no-war scheme $(A^*, W^* \equiv 0)$ is point-wise equivalent to (A, W) if and only if $A^*(\theta) \equiv A(\theta) + W(\theta)$.*

Proof. The proof of necessity follows directly from the definition of point-wise equivalence. If the no-war scheme $(A^*, W^* \equiv 0)$ is point-wise equivalent to (A, W), then $M(\theta; A^*) = M(\theta; A)$ and thus $R(\theta, \theta; A^*) = R(\theta, \theta; A)$. Using (4), it then follows that $A^*(\theta) = A(\theta) + W(\theta)$.

For the proof of sufficiency, we assume that (A, W) satisfies (On-IC) and define A^* by $A^*(\theta) \equiv A(\theta) + W(\theta)$. We must show that $(A^*, W^* \equiv 0)$ satisfies (On-IC) and preserves the original market shares and interim-stage profit under (A, W). Observe first that, if $(A^*, W^* \equiv 0)$ preserves the original market shares, $M(\theta; A^*) = M(\theta; A)$ for all θ, then it satisfies (On-IC) in the Relaxed Program and preserves the original interim-stage profit in (4). Hence, it suffices to show that $(A^*, W^* \equiv 0)$ preserves the original market shares under (A, W). To prove this part, 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 $(A^*, W^* \equiv 0)$ engages in sorting (pooling) are consistent with the intervals on which (A, W) engages in sorting (pooling), and we also show that (A, W) and $(A^*, W^* \equiv 0)$ jump at the same points.

First, suppose that (A, W) entails sorting for $\theta \in [\theta_1, \theta_2] \subset [\underline{\theta}, \overline{\theta}]$. Using (3), we find the interim profit for $\theta \in [\theta_1, \theta_2]$:

$$R(\theta, \theta; A) - A(\theta) - W(\theta) = R(\theta_2, \theta_2; A) - A(\theta_2) - W(\theta_2) + \int_{\theta}^{\theta_2} D(p(x))M(x; A)dx \tag{5}$$

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

$$A(\theta) + W(\theta) - [A(\theta_2) + W(\theta_2)] = -\int_{\theta}^{\theta_2} r(p(x), x)[\partial M(x; A)/\partial x]dx, \tag{6}$$

where $\frac{\partial M(x;A)}{\partial x} = -(N-1)[1 - F(x)]^{N-2} f(x) I < 0$ for all $x < \overline{\theta}$. Using (6), we see that $A^*(\theta)$ defined by $A^*(\theta) \equiv A(\theta) + W(\theta)$ is strictly decreasing and thus entails sorting over $\theta \in [\theta_1, \theta_2]$. Hence, A^* preserves the original market shares for $\theta \in [\theta_1, \theta_2]$. Second, suppose that (A, W) entails pooling for $\theta \in [\theta_1, \theta_2]$. In this case, we may rewrite (5) as $A(\theta) + W(\theta) = A(\theta_2) + W(\theta_2)$. Thus, for $\theta \in [\theta_1, \theta_2]$, A^* entails pooling and preserves the original market share. Third, suppose that (A, W) involves a jump of market-share allocation at a point $\widehat{\theta} \in [\underline{\theta}, \overline{\theta}]$ such that

$$M(\widehat{\theta}; A) > \limsup_{\theta > \widehat{\theta}} M(\theta; A) \equiv M_+(\widehat{\theta}; A). \tag{7}$$

The associated limits from the right for wars and advertising are denoted by $W_+(\widehat{\theta})$ and $A_+(\widehat{\theta})$, respectively. The described jump of market-share allocation at $\widehat{\theta}$ means that $A(\widehat{\theta}) > A_+(\widehat{\theta})$. The level of jump is determined such that the on-schedule constraint is binding at $\widehat{\theta}$:

$$A(\widehat{\theta}) + W(\widehat{\theta}) - [A_+(\widehat{\theta}) + W_+(\widehat{\theta})] = r(p(\widehat{\theta}), \widehat{\theta})[M(\widehat{\theta}; A) - M_+(\widehat{\theta}; A)]. \tag{8}$$

Thus, by (7) and (8), A^* entails a jump at $\widehat{\theta}$, with $A^*(\widehat{\theta}) > A^*_+(\widehat{\theta})$.[21] Given that A^* preserves the original market shares in pooling or sorting intervals, the level of jump under A^* is made such that the on-schedule constraint is binding at $\widehat{\theta}$. ■

Proposition 3 identifies a substitutability between current advertising expenditures and future advertising wars. When a scheme (A, W) satisfies (On-IC) and requires a war ($W(\theta) > 0$ for some

[21] In general, 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 $A^*(\theta) \equiv A(\theta) + W(\theta)$ is nonincreasing.

θ), then we may understand that the expected future payoff is reduced due to the possibility of a future advertising war. Proposition 3 indicates that we may then construct a point-wise equivalent scheme, $(A^*, W^* \equiv 0)$ with $A^*(\theta) \equiv A(\theta)+W(\theta)$, in which the possibility of a future advertising war is eliminated ($W^* \equiv 0$) and current advertising expenditures are increased accordingly ($A^*(\theta) \equiv A(\theta) + W(\theta)$). A war is redundant in this sense.

Together, Propositions 2 and 3 greatly simplify our analysis of the repeated advertising game. Proposition 3 implies that, for any (A, W) that solves the Relaxed Program, there exists a point-wise equivalent no-war scheme, $(A^*, W^* \equiv 0)$ with $A^*(\theta) \equiv A(\theta) + W(\theta)$, that also solves the Relaxed Program. By Proposition 2, 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.[22]. Proposition 2 also provides a program that may be solved in order to characterize A^*.

Our next step is to write the static program identified in Proposition 2 in a more useful form. In particular, using Lemma 1 and $W^* \equiv 0$, we may integrate by parts and state the program that A^* must solve as follows:

No-War Program: The program chooses A to maximize

$$E_\theta \left[R(\theta, \theta; A) - A(\theta)\right] = R(\overline{\theta}, \overline{\theta}; A) - A(\overline{\theta}) + E_\theta \left[D(p(\theta))\frac{F}{f}(\theta)M(\theta; A)\right] \tag{9}$$

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

Notice that expected profit is characterized in the No-War Program in terms of two components. Specifically, we may understand the RHS of (9) as being comprised of the "profit at the top" (i.e., the current-period profit earned by a firm with cost type $\overline{\theta}$) and the expected information rents.[23]

Based on our discussion to this point, we may now establish the following proposition:

Proposition 4. *(i) Suppose (A, W) solves the Relaxed Program, and define A^* by $A^*(\theta) \equiv A(\theta) + W(\theta)$. Then $(A^*, W^* \equiv 0)$ solves the Relaxed Program, and so A^* solves the No-War Program. (ii) If A^* solves the No-War Program and $(A^*, W^* \equiv 0)$ satisfies (On-IC), 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.*

[22] The arguments developed here may also be applied to the class of SPPE in which advertising entails full sorting over $[\underline{\theta}, \overline{\theta}]$ in all periods. In particular, an optimal SPPE within the full sorting class is the stationary (no-war) SPPE in which firms use the advertising equilibrium of the static game in all periods. Thus, for firms to improve on the advertising equilibrium of the static game, they must use an advertising scheme that entails some pooling.

[23] In comparison to the static program identifed in Proposition 2, the No-War Program allows for a larger feasible set of advertising functions. This is because the No-War Program uses (3) to re-state the objective function but does not separately use (3) to restrict the feasible set. Accordingly, in some cases, the No-War Program may admit a solution $A(\theta)$ such that $(A(\theta), W(\theta) \equiv 0)$ does not satisfy (On-IC). In our model of advertising, however, for any solution to the No-War Program that does not satisfy (On-IC), we can deliver the same profit at the top and that same market share allocation (and thus the same expected information rents) with another solution $A^*(\theta)$ to the No-War Program such that $(A^*(\theta), W^*(\theta) \equiv 0)$ does satisfy (On-IC). Hence, one of the solutions to the No-War Program is a solution to the static program in Proposition 2. In our analysis of solutions to the No-War Program below, we are careful to focus on solutions that satisfy the on-schedule incentive constraint and thus that also solve the static program in Proposition 2.

We note that all of our findings to this point are quite general, in that they hold for any demand function D and also for any distribution function F. Further restrictions are required below, however, in order to characterize the advertising schedule A^* that solves the No-War Program.

4.2 Optimal SPPE: Pooling at Zero Advertising

In this subsection, we characterize A^* that solves the No-War Program. We encounter a related problem in the comparison between the advertising and random equilibria in the static model. Bagwell and Lee (2009) 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-War Program.

Proposition 5. *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 at zero advertising following all equilibrium-path histories.*

Proof. Using part (ii) of Proposition 4, we must show that $A^* \equiv 0$ solves the No-War Program, if F is log-concave and demand is sufficiently inelastic or if $\overline{\theta} - \underline{\theta}$ is sufficiently small. Demonstration of this result is sufficient, since $(A^* \equiv 0, W^* \equiv 0)$ clearly satisfies (On-IC). Let A denote any other nonincreasing scheme. Note that $M(\theta; A)$ is then nonincreasing, and recall that $M(\theta; A^*) \equiv \frac{1}{N}$. Consider first the profit at the top term in (9). If A entails any sorting, then $M(\overline{\theta}; A^*) = \frac{1}{N} > M(\overline{\theta}; A)$ and $A^*(\overline{\theta}) = 0 \leq A(\overline{\theta})$. Alternatively, if A is a pooling scheme (at some positive level of advertising), then $M(\overline{\theta}; A^*) = \frac{1}{N} = M(\overline{\theta}; A)$ and $A^*(\overline{\theta}) = 0 < A(\overline{\theta})$. In either case, the profit at the top is strictly higher under A^* than under A. Consider second the expected information rents term in (9). For the special case in which $\overline{\theta} - \underline{\theta}$ approaches zero, expected information rents converge to zero; thus, the profit at the top term dominates if the support of possible cost types is sufficiently small. For the general case in which the support may be large, we define the distribution function $G(\theta; A)$ under A:

$$G(\theta; A) \equiv \frac{\int_{\underline{\theta}}^{\theta} M(x; A) f(x) dx}{\int_{\underline{\theta}}^{\overline{\theta}} M(x; A) f(x) dx}. \tag{10}$$

The distribution function $G(\theta; A^*)$ is similarly defined.[24] The denominators of $G(\theta; A)$ and $G(\theta; A^*)$ represent the (ex ante) expected market share, which equals $\frac{1}{N}$. Since $M(\theta; A^*) = \frac{1}{N}$ crosses $M(\theta; A)$ from below, $G(\theta; A^*)$ first-order stochastically dominates $G(\theta; A)$: $G(\theta; A^*) \leq G(\theta; A)$.[25]

[24] With our definition of the distribution function and analysis of expected information rents for the general case in which the support may be large, we build on arguments made by Athey et al (2004) in their analysis of price collusion.

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

Thus, if $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing, then

$$\int_{\underline{\theta}}^{\overline{\theta}} D(p(\theta))\frac{F}{f}(\theta)dG\left(\theta; A^{*}\right) \geq \int_{\underline{\theta}}^{\overline{\theta}} D(p(\theta))\frac{F}{f}(\theta)dG\left(\theta; A\right). \tag{11}$$

This inequality can be rewritten as

$$E_{\theta}\left[D(p(\theta))\frac{F}{f}(\theta)M(\theta; A^{*})\right] \geq E_{\theta}\left[D(p(\theta))\frac{F}{f}(\theta)M(\theta; A)\right]. \tag{12}$$

Thus, if $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing, then expected information rents are weakly higher under A^{*} than under A. The term $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing when F is log-concave ($\frac{F}{f}(\theta)$ is nondecreasing in θ) and demand is sufficiently inelastic.■

Proposition 5 establishes conditions under which an optimal SPPE exists, wherein firms pool at zero advertising in all periods. As indicated in part (ii) of Proposition 4, the key step is to establish conditions under which the No-War Program is solved with an advertising schedule that entails pooling at zero advertising.[26] Profit at the top is uniquely maximized when firms pool at zero advertising. The maximization of expected information rents, however, is more subtle. When $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing, expected information rents are higher when market share is taken from lower types and redistributed to higher types. Since an incentive-compatible market share allocation function must be nonincreasing, expected information rents are then maximized when the advertising schedule entails pooling, so that the market share allocation function is constant at $\frac{1}{N}$. But whether or not $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing depends on the resolution of conflicting forces. On the one hand, if F is log-concave, then $\frac{F}{f}(\theta)$ is increasing in θ.[27] On the other hand, when demand is downward sloping, $D(p(\theta))$ is decreasing in θ. Thus, if F 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.[28] Additionally, in the special case in which the support of possible cost types is sufficiently small, the No-War Program is solved under pooling at zero advertising, since then the expected information rents can be made sufficiently small that their sign is immaterial.

We now summarize our findings. Propositions 2-4 confirm at a general level that equilibrium-path wars are not necessary for optimality: for any D and F, any optimal SPPE payoff can be achieved by an optimal SPPE that is stationary, wherein firms use A^{*} for all equilibrium histories. We also characterize an optimal SPPE that is stationary. Proposition 5 reports conditions under which an optimal SPPE that is stationary entails $A^{*} \equiv 0$ for all equilibrium histories. Building on these findings, we now show that, under the conditions stated in Proposition 5, *any* optimal SPPE is stationary and entails $A^{*} \equiv 0$ for all equilibrium histories. In this way, we establish the

[26] This step is sufficient, since (On-IC) clearly holds for a pooling, no-war scheme, $(A^{*} \equiv 0, W^{*} \equiv 0)$.

[27] The assumption of log-concavity of F is common in the contract literature and is satisfied by many distribution functions.

[28] We discuss these conflicting considerations in greater detail in our companion paper (Bagwell and Lee, 2009).

uniqueness of the optimal SPPE presented in Proposition 5.

Proposition 6. *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 any optimal SPPE is stationary, wherein firms pool at zero advertising following all equilibrium-path histories.*

Proof. Fix an SPPE in which firms do *not* pool at zero following all equilibrium path histories. We may translate the factorization of this SPPE into a scheme (A, W) that satisfies the constraints of the Relaxed Program. Using Proposition 3, the scheme (A, W) is point-wise equivalent to a no-war scheme $(A^*, W^* \equiv 0)$, where $A^* \equiv A + W$. Given the assumed properties of the SPPE, A^* is not identically zero. As established in the proof of Proposition 5, if F is log-concave and demand is sufficiently inelastic, or if the support of possible cost types is sufficiently small, then the No-War Program is uniquely solved when advertising is identically equal to zero. Under these conditions, therefore, the posited no-war scheme $(A^*, W^* \equiv 0)$ can be strictly improved upon by an alternative no-war scheme in which firms *do* pool at zero advertising. Further, for δ sufficiently high, we know from Proposition 4 that the alternative scheme corresponds to an optimal SPPE that exists, is stationary and entails firms pooling at zero advertising following all equilibrium-path histories. Thus, under the stated conditions, a stationary SPPE exists with the described properties which generates a strictly higher expected ex ante profit for firms than does any other SPPE.■

While we allow for a wide range of SPPE advertising behaviors in the repeated game, we show that important conditions exist under which advertising behavior in any optimal SPPE takes a remarkably simple form: along the equilibrium path, no firm advertises in any period. Intuitively, the conditions in the proposition favor pooling; and wars are thus redundant in this context, since the associated payoffs can be achieved by pooling at a higher level of advertising in the current period. Furthermore, pooling at a positive level of advertising in the current period is a wasteful means for firms of achieving the associated market share allocation. They can achieve the same allocation more profitably by pooling at zero advertising.

Propositions 5 and 6 thus provide a formal confirmation of the idea that, even if advertising is legal and informed consumers are responsive to it, firms may eliminate advertising as part of an optimal self-enforcing collusive agreement. When firms collude in this way, the welfare of consumers is reduced from the welfare that they enjoy in the non-cooperative advertising equilibrium. Given that the induced distribution of monopoly prices is not altered, uninformed consumers' surplus remains unaffected. The collusive agreement, however, prevents informed consumers from using advertising to infer the identity of the lowest-cost, and thus the lowest-price, firm in the market. The average transaction price is thus higher when advertising is eliminated as part of a collusive agreement among firms.

We emphasize that Propositions 5 and 6 may hold even when demand is elastic. First, observe that these propositions hold for any demand function, if the support of possible cost types is sufficiently small and δ is sufficiently high. Second, consider the constant-elasticity demand function,

$D(p) = p^{-\epsilon}$, and suppose that demand is elastic (i.e., $\epsilon > 1$). If θ is distributed uniformly over $[\underline{\theta}, \overline{\theta}]$ where $\underline{\theta} > 0$, then $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing when $\overline{\theta}/[\overline{\theta} - \underline{\theta}] > \epsilon$; thus, any optimal SPPE for patient firms entails pooling at zero advertising, 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.

4.3 Optimal SPPE: Partial Pooling

While Proposition 6 isolates an important set of conditions under which any optimal SPPE takes a very simple form, it is also interesting to consider the form that optimal SPPE may take when these conditions fail. In this subsection, without requiring that $D(p(\theta))\frac{F}{f}(\theta)$ is everywhere nondecreasing, we establish that optimal SPPE for patient firms involves at least partial pooling.[29]

A difficulty with solving the No-War 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 $[\underline{\theta}, y]$. This nonincreasing scheme has the following expected profit:[30]

$$\begin{aligned} E_\theta\left[R(\theta, \theta; A) - A(\theta)\right] &= r\left(p(\overline{\theta}), \overline{\theta}\right) M(\overline{\theta}; A) + \int_{\underline{\theta}}^{y} D(p(\theta))\frac{F}{f}(\theta) M\left(\theta; A\right) dF(\theta) \\ &\quad + \int_{y}^{\overline{\theta}} D(p(\theta))\frac{F}{f}(\theta) M\left(\theta; A\right) dF(\theta). \end{aligned} \tag{13}$$

The market share allocation functions are given by $M(\theta; A) = \frac{U}{N} + [1 - F(\theta)]^{N-1} I$ for $\theta \in [\underline{\theta}, y]$ and $M(\theta; A) = \frac{U}{N} + [1 - F(y)]^{N-1}\frac{I}{N}$ for $\theta \in (y, \overline{\theta}]$. The level of jump is determined such that the on-schedule constraint is binding at y:

$$A(y) = r(p(y), y)[1 - F(y)]^{N-1} I \left(1 - \frac{1}{N}\right). \tag{14}$$

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 at least a small pooling interval at the top, $(y, \overline{\theta}]$.

We extend this result and develop two general points. First, any no-war scheme that has a sorting interval at the top can be improved upon by an alternative no-war scheme that has a pooling interval at the top (i.e., an interval $(y, \overline{\theta}]$ on which $A(\theta) = 0$).[31] Second, if firms are

[29] As above, we solve the No-War Program to characterize optimal SPPE. If the solution to the program involves positive advertising as in partial pooling, then optimal SPPE may take the form of a stationary or non-stationary equilibrium; the reason is that firms may then allocate advertising expenses across periods, because of the substitutability between current-period advertising and future advertising wars. If firms implement a non-stationary SPPE, then they move between cooperative and war phases in their advertising conduct.

[30] The expression for expected profit is derived in the Appendix.

[31] The proof for this part is provided by the proof of Lemma 2 in the Appendix.

sufficiently patient, then any optimal SPPE entails a pooling interval at the top. For the proof of this second part, assume that an optimal SPPE exists that entails a sorting interval at the top. We may translate the associated factorization into a scheme (A, W) that satisfies the constraints of the Relaxed Program. From here, we can construct a point-wise equivalent scheme $(A^*, W^* \equiv 0)$. Since this no-war scheme also has a sorting interval at the top, it can be improved upon by an alternative no-war scheme that has a pooling interval at the top. For sufficiently patient firms, we can then support an SPPE in which the advertising schedule from the alternative no-war scheme is used in each period along the equilibrium path. Our inital assumption is thus contradicted.

Lemma 2. *For any F and D, if δ is sufficiently high, then any optimal SPPE has a pooling interval $(y, \overline{\theta}]$ on which $A(\theta) = 0$.*

To present a more comprehensive characterization of optimal SPPE, we next assume that the entire advertising schedule A is represented by K finite 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 $\theta_k < \theta_{k+1}$. Referring to Proposition 4 and Lemma 2, we now restrict attention to stationary (no-war) SPPE which entail pooling at zero advertising on an interval at the top; straightforward arguments as above ensure that the findings below hold for any optimal SPPE for patient firms. If the schedule A solves the No-War Program, then expected profit is

$$\begin{aligned} E_\theta \left[R\left(\theta, \theta; A\right) - A\left(\theta\right) \right] = {} & r\left(p(\overline{\theta}), \overline{\theta}\right) M(\overline{\theta}; A) \\ & + \sum_{k=1}^{K} \int_{\theta_k}^{\theta_{k+1}} D(p(\theta)) \frac{F}{f}(\theta) M(\theta; A) dF(\theta). \end{aligned} \tag{15}$$

The market share for $\theta \in (\theta_K, \theta_{K+1}]$ is $M(\theta; A) = \frac{U}{N} + [1 - F(\theta_K)]^{N-1} \frac{I}{N}$. If $(\theta_k, \theta_{k+1}]$ is a pooling interval, then, for $\theta \in (\theta_k, \theta_{k+1}]$,

$$M(\theta; A) = \frac{U}{N} + \sum_{j=0}^{N-1} \binom{N-1}{j} \frac{1}{j+1} \left[F(\theta_{k+1}) - F(\theta_k) \right]^j \left[1 - F(\theta_{k+1}) \right]^{N-j-1} I. \tag{16}$$

If $(\theta_k, \theta_{k+1}]$ is a sorting interval, then, for $\theta \in (\theta_k, \theta_{k+1}]$, $M(\theta; A) = \frac{U}{N} + \left[1 - F(\theta)\right]^{N-1} I$. The expected market share over the entire interval is $\frac{1}{N}$:

$$\sum_{k=1}^{K} \int_{\theta_k}^{\theta_{k+1}} M(\theta; A) f(\theta) d\theta = \frac{1}{N}. \tag{17}$$

An advertising schedule has a discontinuity (a jump) between two pooling intervals and between sorting and pooling intervals. The level of jump at a point is determined by the binding (On-IC) at that point.

We next show that optimal SPPE reflect forces in favor of pooling in a range of cost types where $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing.[32] Suppose that a scheme A is sorting *and* $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing in a range $(\theta_i, \theta_{i+1}]$. As we show in detail in the Appendix, we can then construct an alternative scheme A^* such that $A^*(\theta)$ preserves the original scheme $A(\theta)$ for $\theta > \theta_{i+1}$ and $A^*(\theta)$ is pooling for $\theta \in (\theta_i, \theta_{i+1}]$ and makes a parallel shift from $A(\theta)$ for $\theta \leq \theta_i$.[33] Given that $A^*(\theta)$ replaces sorting with pooling for $\theta \in (\theta_i, \theta_{i+1}]$, the original market shares under A are affected by A^* for $\theta \in (\theta_i, \theta_{i+1}]$. For the affected range $(\theta_i, \theta_{i+1}]$, we define the distribution function under A^*:

$$G(\theta_i, \theta_{i+1}; A^*) \equiv \frac{\int_{\theta_i}^{\theta} M(x; A^*) f(x) dx}{\int_{\theta_i}^{\theta_{i+1}} M(x; A^*) f(x) dx} \text{ for } \theta \in (\theta_i, \theta_{i+1}]. \tag{18}$$

The distribution $G(\theta_i, \theta_{i+1}; A)$ is analogously defined under A. We confirm in the Appendix that the denominators of the two distribution functions are the same. Since $M(\theta; A^*)$ crosses $M(\theta; A)$ from below in the range $(\theta_i, \theta_{i+1}]$, $G(\theta_i, \theta_{i+1}; A^*)$ first-order stochastically dominates $G(\theta_i, \theta_{i+1}; A)$. We may invoke the argument used in the proof of Proposition 5 and compare the information-rent terms:

$$\int_{\theta_i}^{\theta_{i+1}} D(p(\theta))\frac{F}{f}(\theta) M(\theta; A^*) dF(\theta) \geq \int_{\theta_i}^{\theta_{i+1}} D(p(\theta))\frac{F}{f}(\theta) M(\theta; A) dF(\theta). \tag{19}$$

Since A^* is designed to preserve the original market share under A other than in the range $(\theta_i, \theta_{i+1}]$, the expected profit remains the same except for the information-rent terms in (19). Hence, we conclude that the expected profit is weakly higher under A^* than under A.

This finding can be readily extended. Suppose that $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing for $\theta \in (\theta_i, \theta_{i+1}]$ and is *strictly* increasing for some interior type $\theta \in (\theta_i, \theta_{i+1})$. In this case, expected profit is strictly higher under A^* than under A, from which it follows that no optimal SPPE is sorting for $\theta \in (\theta_i, \theta_{i+1}]$.[34] Likewise, we may establish that any optimal SPPE involves pooling at the bottom (i.e., for a range $[\underline{\theta}, x]$ where $x \leq \overline{\theta}$). This is because, for any F and D, $D(p(\theta))\frac{F}{f}(\theta)$ strictly increases at the neighborhood of $\underline{\theta}$, given $f(\underline{\theta}) > 0$.

We summarize our findings as follows:

Proposition 7. *Assume that δ is sufficiently high. (i) For any F and D, any optimal SPPE involves pooling at the bottom and at the top. (ii) If $D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing for $\theta \in (\theta_i, \theta_{i+1}]$ and is strictly increasing for some interior type in this range, then no optimal SPPE entails sorting for $\theta \in (\theta_i, \theta_{i+1}]$.*

The proof is in the Appendix. Proposition 7 has two implications. First, since the repeated play of the advertising equilibrium of the static game is a stationary SPPE that entails full sorting, Proposition 7 ensures that any optimal SPPE for patient firms involves at least partial pooling and strictly improves upon the repeated use of the advertising equilibrium. Second, if a sorting interval

[32] Our analysis refers to two related intervals: (i) the interval of θ on which $A(\theta)$ is defined and (ii) the interval of θ on which $D(p(\theta))\frac{F}{f}(\theta)$ is defined. To avoid confusion, we hereafter refer to the latter interval as the "range" of θ.

[33] The definition of A^* and associated proofs are detailed in the proof of Proposition 7 in the Appendix.

[34] In other words, for this case, in an optimal SPPE, it cannot be the true that, for all $\theta \in (\theta_i, \theta_{i+1}]$, $A(\theta)$ is strictly decreasing.

is ever used by patient firms, then it is restricted to an "intermediate range" in which $D(p(\theta))\frac{F}{f}(\theta)$ is nonincreasing.

4.4 Optimal SPPE: At Most Two Pooling Intervals

In Proposition 7 of the previous subsection, we show that optimal SPPE exhibit robust forces in favor of at least partial pooling. In this subsection, we go further and establish that optimal SPPE may use quite a small number of pooling steps and restrict the use of sorting interval as a means of reducing advertising expenses. To develop these points, we now restrict attention to stationary (no-war) SPPE that entail pooling at the top and at the bottom. We note, though, that our findings presented below hold for any optimal SPPE. Our analysis is founded on the following assumption on F and D.

Assumption 1. *$D(p(\theta))\frac{F}{f}(\theta)$ is strictly quasiconcave with a maximizer $\theta^* \in (\underline{\theta}, \overline{\theta}]$.*

Assumption 1 holds if $D(p(\theta))\frac{F}{f}(\theta)$ has a unique interior maximizer $\theta^* \in (\underline{\theta}, \overline{\theta})$ and is strictly increasing for $\theta < \theta^*$ and strictly decreasing for $\theta > \theta^*$. It also includes the case where $D(p(\theta))\frac{F}{f}(\theta)$ is everywhere strictly increasing with a maximizer, $\overline{\theta}$. Given that $D(p(\theta))$ is strictly decreasing in θ, the assumption is satisfied in a wide range of settings when F is log-concave. For example, Assumption 1 holds if F is the uniform distribution on $[\underline{\theta}, \overline{\theta}]$ and $D(p) = (\alpha - \beta p)^\gamma$, where $\underline{\theta} \geq 0$, $(\alpha, \beta, \gamma) > 0$ and $\alpha - \beta\overline{\theta} > 0$. Notice that $D(p)$ is linear when $\gamma = 1$ and is convex (concave) when $\gamma > 1$ ($\gamma < 1$). Similarly, Assumption 1 holds if F is the uniform distribution on $[\underline{\theta}, \overline{\theta}]$ and $D(p) = p^{-\epsilon}$, where $\underline{\theta} > 0$ and $\epsilon > 1$. Using the two demand functions just presented, we can also numerically confirm that Assumption 1 is satisfied for a substantially wide range of parameters when F is a "truncated" normal distribution.[35] Assumption 1 may be violated in cases where demand is very convex and cost types are distributed with low variance.[36]

We now show that, under Assumption 1, any optimal SPPE for patient firms entails *at most two* pooling intervals. For notational simplicity, let $\phi(\theta) \equiv D(p(\theta))\frac{F}{f}(\theta)$. We focus on the case in which $\phi(\theta)$ is strictly quasiconcave with an interior maximizer $\theta^* \in (\underline{\theta}, \overline{\theta})$, relegating related proofs to the Appendix. We present our findings in two steps. First, an optimal SPPE cannot have two *separate* pooling steps *within* a range where $\phi'(\theta) > 0$. If a scheme includes two separate pooling intervals, $(\theta_i, y]$ and $(y, \theta_{i+1}]$, within a range $(\theta_i, \theta_{i+1}]$ where $\phi'(\theta) > 0$, then there is an alternative scheme that replaces the two pooling steps with one pooling step for $\theta \in (\theta_i, \theta_{i+1}]$. Using the distributions for $\theta \in (\theta_i, \theta_{i+1}]$ as in (18), we find that the expected profit is strictly higher under the alternative scheme. By the same token, other than the pooling interval at the top, an optimal SPPE cannot

[35] The normal distribution with mean and variance, μ and σ^2, has density $\lambda(x) \equiv \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$ where $-\infty < x < \infty$. The distribution function is $\Lambda(x) = \int_{-\infty}^{x} \lambda(t)dt$. The density under a truncated normal distribution is defined as $f(\theta) = \frac{\lambda(\theta)}{\Lambda(\overline{\theta}) - \Lambda(\underline{\theta})}$ if $\underline{\theta} \leq \theta \leq \overline{\theta}$, and $f(\theta) = 0$ otherwise. The associated distribution function is $F(\theta) = \int_{\underline{\theta}}^{\theta} f(x)dx$.

[36] We have numerically confirmed that the assumption is violated only in a very limited range of parameters, as in the case where $D(p) = (\alpha - \beta p)^\gamma$ is convex and F is normal with a very low variance (e.g., $D(p(\theta))\frac{F}{f}(\theta)$ has two local maximizers when $\underline{\theta} = 0, \overline{\theta} \approx 0.5$, $\alpha = 1$, $\beta = 2$, $\gamma = 2$, $\mu = 0.25$ and $\sigma^2 = 0.01$). Note that we set $\alpha = 1$, to be consistent with the assumption of unit mass of consumers.

include a separate pooling interval *within* a range $(\theta_i, \theta_{i+1}]$ where $\phi'(\theta) < 0$. If a scheme has a pooling interval $(y, z]$ within $(\theta_i, \theta_{i+1}]$ where $\phi'(\theta) < 0$, then there is an alternative scheme that replaces pooling with sorting for $\theta \in (y, z]$. We also find that the expected profit is strictly higher under the alternative scheme.

Second, an optimal SPPE cannot have three pooling steps; equivalently, it cannot have an *intermediate* pooling interval.[37] Remember that an optimal SPPE cannot include two separate pooling steps or any sorting interval within $[\underline{\theta}, \theta^*]$ where $\phi'(\theta) > 0$, and that, other than at the top, an optimal SPPE cannot have a separate pooling step within $(\theta^*, \overline{\theta}]$ where $\phi'(\theta) < 0$. Thus, the one remaining possibility for an optimal SPPE candidate A to have three pooling steps is that the scheme has pooling steps, $[\underline{\theta}, y]$, $(y, z]$ and $(z, \overline{\theta}]$, such that $\underline{\theta} < y < \theta^* < z < \overline{\theta}$.[38] Consider an alternative scheme A^* that has pooling steps, $[\underline{\theta}, y^*]$, $(y^*, z]$ and $(z, \overline{\theta}]$, such that $\underline{\theta} < \theta^* < y^* < z < \overline{\theta}$. The alternative scheme lengthens the first pooling step beyond the range $[\underline{\theta}, \theta^*)$ where $\phi'(\theta) > 0$. The original market shares under A are affected by A^* for $\theta \in [\underline{\theta}, z]$. Given the affected range $[\underline{\theta}, z]$, the market shares for types at the bottom $[\underline{\theta}, y]$ and at the top $(y^*, z]$ are lower under A^* than under A, while the market shares for types in the intermediate range $(y, y^*]$ are higher under A^* than under A. We define the distribution function under A^*:

$$G(\underline{\theta}, z; A^*) \equiv \frac{\int_{\underline{\theta}}^{\theta} M(x; A^*) f(x) dx}{\int_{\underline{\theta}}^{z} M(x; A^*) f(x) dx} \text{ for } \theta \in [\underline{\theta}, z]. \tag{20}$$

The distribution $G(\underline{\theta}, z; A)$ is similarly defined under A. The two functions have the same denominators. Differentiation of the functions with respect to θ shows that, given the range $[\underline{\theta}, z]$, the slope of $G(\underline{\theta}, z; A^*)$ is flatter (steeper) than that of $G(\underline{\theta}, z; A)$ at the bottom $[\underline{\theta}, y]$ and at the top $(y^*, z]$ (in the intermediate range $(y, y^*]$). The point y^* is chosen such that $G(\underline{\theta}, z; A^*)$ crosses $G(\underline{\theta}, z; A)$ *from below* at the point θ^*.[39] Letting $\Delta(\theta) \equiv G(\underline{\theta}, z; A) - G(\underline{\theta}, z; A^*)$, we then compare the affected information-rent terms:

$$\begin{aligned} &\int_{\underline{\theta}}^{z} \phi(\theta) dG(\underline{\theta}, z; A^*) - \int_{\underline{\theta}}^{z} \phi(\theta) dG(\underline{\theta}, z; A) \\ &= \int_{\underline{\theta}}^{\theta^*} \phi'(\theta) \Delta(\theta) d\theta + \int_{\theta^*}^{z} \phi'(\theta) \Delta(\theta) d\theta > 0. \end{aligned} \tag{21}$$

We thus find that the expected profit is strictly higher under A^* than under A. Intuitively, A^* is more suitable than A to increase the expected information rents: in the range where $\phi'(\theta) > 0$, A^* decreases the market share for types below $y \in (\underline{\theta}, \theta^*)$ and increases the market share for types above y; and in the range where $\phi'(\theta) < 0$, A^* increases the market share for types below $y^* \in (\theta^*, z)$

[37] From the analysis below, it follows that an optimal SPPE cannot have more than three pooling steps.

[38] The scheme A may include a sorting interval between $(y, z]$ and $(z, \overline{\theta}]$ where $\phi'(\theta) < 0$. The findings below are not affected by this change.

[39] Observe that $y^* \in (\theta^*, z)$. If $y^* \leq \theta^*$, then $G(\underline{\theta}, z; A^*)$ crosses $G(\underline{\theta}, z; A)$ at type $\theta < \theta^*$, since the slope of $G(\underline{\theta}, z; A^*)$ is flatter than that of $G(\underline{\theta}, z; A)$ in the range $(y^*, z]$. If $y^* = z$, then $G(\underline{\theta}, z; A^*)$ crosses $G(\underline{\theta}, z; A)$ from below at the endpoint $z > \theta^*$.

and decreases the market share for types above y^*. Observe further that, other than the pooling at the top $(z, \overline{\theta}]$, A^* includes a separate pooling step $(y^*, z]$ within the range $(\theta^*, \overline{\theta}]$ where $\phi'(\theta) < 0$. We can then construct another alternative scheme that replaces pooling with sorting for $\theta \in (y^*, z]$ and strictly improves upon the scheme A^*. Hence, an optimal SPPE cannot have an intermediate pooling interval.

We now summarize our findings:

Proposition 8. *Suppose that δ is sufficiently high and that Assumption 1 is satisfied. (i) An optimal SPPE has at most two pooling intervals: it is characterized by either (a) one full pooling step at zero advertising or (b) two pooling steps with or without an intermediate sorting interval. (ii) If an optimal SPPE ever includes an intermediate sorting interval, then it restricts the sorting interval to a subset of the range in which $D(p(\theta))\frac{F}{f}(\theta)$ is strictly decreasing.*

The proof is in the Appendix. For δ sufficiently high, Proposition 8 complements the finding that an optimal SPPE must entail pooling at the bottom and at the top. Under Assumption 1, if an optimal SPPE ever involves sorting, it uses sorting *only once* in an intermediate range of cost types; hence, the possible forms of optimal SPPE are characterized by the two cases specified in Proposition 8 (i). Whether an optimal SPPE includes a sorting scheme depends on the extent to which $D(p(\theta))\frac{F}{f}(\theta)$ decreases in an intermediate range and the magnitude of the profit at the top.[40] Proposition 8 thus confirms the idea that, under general demand functions, patient firms may use a small number of pooling steps and restrict the use of sorting intervals as a means of reducing the intensity of advertising competition.[41] To the detriment of consumer welfare, such a collusive agreement restricts the informed consumers' capacity to locate the lowest available price in the market.

5 Off-Schedule Incentive Constraints

Up to this point, we have ignored off-schedule incentive constraints by assuming that firms are sufficiently patient. We now consider off-schedule constraints and characterize the critical discount factor above which they are satisfied. In particular, motivated by our findings above, we characterize the critical discount factor for optimal SPPE that are stationary and entail pooling at zero advertising or entail two pooling intervals with an intermediate sorting interval.

We first characterize the critical discount factor, $\widehat{\delta}^p \in (0, 1)$, above which an optimal SPPE exists that is stationary and entails pooling at zero advertising (as established in Proposition 5). When firms pool at zero advertising, a firm faces a temptation to cheat by advertising a small,

[40] Suppose that $D(p) = 1 - p$, $N = 5$ and F is the uniform distribution on $[0, \overline{\theta}]$. Then, $D(p(\theta))\frac{F}{f}(\theta) = \frac{(1-\theta)\theta}{2}$ is concave with a maximizer 0.5. If $\overline{\theta} = 0.99$, then any optimal SPPE has two pooling steps with an intermediate sorting interval approximately on $(0.752, 0.962]$. If $\overline{\theta} = 0.77$, then any optimal SPPE has only two pooling steps with a jump at 0.75. If $\overline{\theta} = 0.70$, then it has a single pooling step at zero advertising.

[41] We can also show that Proposition 8 holds even in settings where Assumption 1 may fail. In particular, Proposition 8 holds as well if $D(p(\theta))\frac{F}{f}(\theta)$ is strictly quasiconcave for $\theta \leq \theta^{**} \in (\underline{\theta}, \overline{\theta})$ with a local maximizer θ^* and is strictly quasiconvex for $\theta > \theta^*$ with a local minimizer $\theta^{**} > \theta^*$.

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.[42] 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:

$$r(p(\underline{\theta}), \underline{\theta}) I \left(1 - \frac{1}{N}\right) \leq \frac{\delta}{1-\delta} \left(\pi^p - \pi^s\right), \tag{22}$$

where $\pi^s \equiv E_\theta \left[\Pi(\theta, \theta; A)\right]$ 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. The expected per-period profit under the advertising equilibrium is directly characterized by (9) in the No-War Program:

$$E_\theta \left[\Pi(\theta, \theta; A)\right] = r(p(\overline{\theta}), \overline{\theta}) \frac{U}{N} + E_\theta \left[D(p(\theta)) \frac{F}{f}(\theta) M(\theta; A)\right] \tag{23}$$

where $M(\theta; A) = \frac{U}{N} + [1 - F(\theta)]^{N-1} I$ and we use that $A(\overline{\theta}) = 0$.

Solving (22) for the critical discount factor, we obtain that pooling at zero advertising satisfies the off-schedule constraint if

$$\delta \geq \widehat{\delta}^p \equiv \frac{r(p(\underline{\theta}), \underline{\theta})(N-1) I}{r(p(\underline{\theta}), \underline{\theta})(N-1) I + N\left(\pi^p - \pi^s\right)}. \tag{28}$$

As shown in Proposition 6, $\pi^p > \pi^s$ if F is log-concave and demand is sufficiently inelastic, or if $\overline{\theta} - \underline{\theta}$ 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*

[42] Other symmetric punishments, such as those that take a "carrot-stick" form, may also be considered. Building on arguments developed by Athey et al. (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.

possible cost types is sufficiently small, then, for all $\delta \geq \widehat{\delta}^p$, there exists an optimal SPPE, and in any optimal SPPE firms pool at zero advertising following all equilibrium-path histories.

In comparison to Proposition 6, Proposition 9 provides an explicit characterization of the critical discount factor above which firms can enforce the unique optimal SPPE outcome, wherein no firm advertises in any period.

We next characterize the critical discount factor, $\widehat{\delta}^{2p} \in (0,1)$, above which an optimal SPPE exists that is stationary and entails two pooling intervals. Suppose that a scheme A^* solves the No-War Program and has two pooling steps, $[\underline{\theta}, y]$ and $(z, \overline{\theta}]$, and an intermediate sorting interval $(y, z]$.[43] The scheme has jumps at y and at z. Any off-schedule deviation, $\widehat{a}$, takes the form of either (i) $\widehat{a} > A^*(\theta)$ for $\theta \in [\underline{\theta}, y]$ or (ii) $\widehat{a} < A^*(\theta)$ for $\theta \in [\underline{\theta}, y]$. The first deviation is to "out-advertise" firms on the interval $[\underline{\theta}, y]$, and the second deviation is to out-advertise firms on the interval $(y, z]$ or $[z, \overline{\theta}]$. We now consider the off-schedule incentive constraints that are associated with these two deviations.

We begin with the first deviation. This deviation captures all informed consumers. We show that a firm with type $\underline{\theta}$ has the greatest short-term gain from the first deviation. Consider first the corresponding off-schedule constraint for $\theta \in [\underline{\theta}, y]$. Among the types on $[\underline{\theta}, y]$, a firm with cost type $\underline{\theta}$ has the greatest short-term gain from the deviation. We may represent the off-schedule constraint for $\underline{\theta}$ as

$$r(p(\underline{\theta}), \underline{\theta}) \left(\frac{U}{N} + I - M(\underline{\theta}; A^*) \right) \leq \frac{\delta}{1-\delta} \left(\pi^{2p} - \pi^s \right), \tag{25}$$

where $\pi^{2p} \equiv E_\theta \left[\Pi(\theta, \theta; A^*) \right]$ is a firm's expected per-period profit under the two-step scheme A^* described above. The LHS represents the short-term gain for $\underline{\theta}$. Consider next the off-schedule constraint for $\theta > y$. The short-term gain for $\theta \in (y, \overline{\theta}]$ is

$$\begin{aligned} & r(p(\theta), \theta) \left(\frac{U}{N} + I - M(\theta; A^*) \right) - \left[A^*(\underline{\theta}) - A^*(\theta) \right] \\ & = r(p(\theta), \theta) \left(\frac{U}{N} + I - M(\underline{\theta}; A^*) \right) - \left[\Pi(\theta, \theta; A^*) - \Pi(\underline{\theta}, \theta; A^*) \right], \end{aligned} \tag{26}$$

where the equality utilizes $\Pi(\underline{\theta}, \theta; A^*) = r(p(\theta), \theta) M(\underline{\theta}; A^*) - A^*(\underline{\theta})$. Note that $\Pi(\theta, \theta; A^*) \geq \Pi(\underline{\theta}, \theta; A^*)$ is ensured by (On-IC). It then follows that the short-term gain for $\underline{\theta}$ in (25) is greater than for $\theta \in (y, \overline{\theta}]$ in (26).

For the first deviation, we may thus solve (25) for the critical discount factor. We find that the first deviation is unattractive to a firm with type $\underline{\theta}$, and thus to a firm with any type θ, if

$$\delta \geq \widehat{\delta}_{\underline{\theta}, y} \equiv \frac{r(p(\underline{\theta}), \underline{\theta})(1 - \mu(y)) I}{r(p(\underline{\theta}), \underline{\theta})(1 - \mu(y)) I + (\pi^{2p} - \pi^s)}. \tag{27}$$

[43] Our analysis can be readily modified to characterize the critical discount factor when the scheme has only two pooling intervals ($y = z$).

Since $[\underline{\theta}, y]$ is a pooling interval, $\mu(y)$ in (27) is defined by

$$\mu(y) \equiv \sum_{j=0}^{N-1} \binom{N-1}{j} \frac{1}{j+1} [F(y)]^j [1-F(y)]^{N-j-1} . \tag{28}$$

Given our assumption that A^* has two pooling regions and solves the No-War Program, we have that $\mu(y) \in (0,1)$ and $\pi^{2p} > \pi^s$, from which it follows that $\widehat{\delta}_{\underline{\theta},y} \in (0,1)$.

We next explore the second deviation. We show that a firm with type z has the greatest short-term gain from the second deviation. Consider first a deviation with $\widehat{a}$ that is slightly above the on-schedule advertising $A^*_+(y)$, where $A^*_+(y) \equiv \limsup_{\theta > y} A^*(\theta)$ represents the limit from the right. Under this deviation, a firm out-advertises firms on the sorting interval $(y, z]$ and thus obtains the market share $\frac{U}{N} + [1-F(y)]^{N-1}I$. Note that any firm with type θ can earn the same market share when it chooses the *on-schedule* advertising $A^*_+(y)$. Hence, as long as (On-IC) holds for θ, then the firm with type θ will not undertake such an off-schedule deviation. Consider next a deviation with $\widehat{a}$ that is slightly above zero. With this deviation, a firm out-advertises firms on $(z, \overline{\theta}]$ and thus obtains the market share $\frac{U}{N} + [1-F(z)]^{N-1}I$. Note that any firm with type θ can earn the same market share when it chooses the *on-schedule* advertising $A^*(z)$. Thus, the short-term gain from the second deviation for type θ becomes

$$\begin{aligned} & r(p(\theta), \theta) [M(z; A^*) - M(\theta; A^*)] + A^*(\theta) \\ & = \Pi(z, \theta; A^*) - \Pi(\theta, \theta; A^*) + A^*(z) \leq A^*(z), \end{aligned} \tag{29}$$

where the inequality follows since (On-IC) ensures that $\Pi(\theta, \theta; A^*) \geq \Pi(z, \theta; A^*)$. The RHS of the inequality, $A^*(z)$, represents the short-term gain when $\theta = z$. Thus, a firm with type z gains the most from out-advertising firms on $(z, \overline{\theta}]$. Observe that a firm with type z's short-term gain, $A^*(z)$, is the level of the jump made at z such that (On-IC) is binding:

$$A^*(z) = r(p(z), z)[1-F(z)]^{N-1} I \left(1 - \frac{1}{N}\right). \tag{30}$$

We may thus represent the off-schedule constraint for type z as

$$r(p(z), z)[1-F(z)]^{N-1} I \left(1 - \frac{1}{N}\right) \leq \frac{\delta}{1-\delta} \left(\pi^{2p} - \pi^s\right). \tag{31}$$

For the second deviation, we may thus solve (31) for the critical discount factor. We find that the second deviation is unattractive to a firm with type z, and thus to a firm with any type θ, if

$$\delta \geq \widehat{\delta}_{z,\overline{\theta}} \equiv \frac{r(p(z), z)(N-1)[1-F(z)]^{N-1} I}{r(p(z), z)(N-1)[1-F(z)]^{N-1} I + N\left(\pi^{2p} - \pi^s\right)}. \tag{32}$$

Arguing as above, we can establish that $\widehat{\delta}_{z,\overline{\theta}} \in (0,1)$.

We are now ready to summarize our findings concerning off-schedule constraints and the posited optimal SPPE with two pooling intervals. In particular, the no-war scheme A^* satisfies all off-schedule constraints if (27) and (32) are satisfied; thus, this scheme satisfies all off-schedule constraints if $\delta \geq \widehat{\delta}^{2p} \equiv \max\{\widehat{\delta}_{\underline{\theta},y}, \widehat{\delta}_{z,\overline{\theta}}\}$, where the critical discount factor, $\widehat{\delta}^{2p}$, satisfies $\widehat{\delta}^{2p} \in (0,1)$.

Thus far, we have characterized critical discount factors within the class of stationary (no-war) SPPE. As we show in previous sections, for sufficiently patient firms, the use of a stationary (no-war) scheme does not limit the scope of optimal SPPE: the payoffs achieved in any optimal non-stationary SPPE can always be achieved as well in an optimal SPPE that is stationary. In fact, this same result holds as well when firms are less patient and off-schedule constraints may bind. Intuitively, if an off-schedule constraints 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, et al (2004) provide a related argument in their analysis of price collusion among impatient firms, and so we do not develop this point in further detail here.

6 Conclusion

We investigate the advertising behavior of firms with private information as to their respective costs. We begin by considering a static advertising game in which each firm's advertising choice may signal its costs, and thus its price, to those consumers who are informed of advertising activities. In the static game, an advertising equilibrium exists, in which informed consumers use an advertising search rule whereby they buy from the highest-advertising firm. In this equilibrium, non-price advertising directs consumers to the firm with the lowest cost and price in the market.[44] We next analyze a repeated advertising game in which privately informed firms may achieve a self-enforcing agreement to limit the use of advertising. We observe that firms face trade-offs when selecting an optimal collusive scheme: while the use of advertising can direct sales to lower-cost firms and thereby promote productive efficiency, it can do so only when sufficient current or future advertising expenses are incurred. If firms sacrifice productive efficiency by pooling at zero advertising, they can eliminate current and future advertising expenses. Allowing for a wide range of collusive advertising behaviors, we establish conditions under which optimal collusion entails pooling at zero advertising. We also show that, under general conditions, optimal collusion involves at least partial pooling and thus strictly improves upon the repeated use of the static advertising equilibrium. In summary, non-price advertising can promote product efficiency and raise consumer welfare; however, colluding firms often have incentive to limit the use of non-price advertising.

We close by mentioning two possible extensions of the model. A first possibility is that advertising by any one firm may have a public-good aspect 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'

[44] The static advertising game is analyzed in greater detail in our companion paper (Bagwell and Lee, 2009).

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. In this setting, colluding firms may have incentive to share advertising expenses. For new-product markets in particular, an analysis of such collusive advertising is an important direction for future work.

A second possibility is to extend our analysis to allow for asymmetric equilibria. In their price-collusion model, Athey and Bagwell (2001) show that profit may be higher in asymmetric perfect public equilibria 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 suffers a reduced market share in the future. Rival firms then enjoy a future market share gain. In this way, asymmetric equilibria allow that continuation values may be used to satisfy on-schedule constraints, without requiring that firms symmetrically experience a reduced continuation value.[45] In their model, consumers observe prices and have no independent interest in firms' costs. By contrast, in our advertising model, informed consumers observe advertising and draw inferences as to costs and thus prices. The construction of asymmetric equilibria may be more challenging in this context. Suppose 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. If informed consumers 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 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.

7 Appendix

This appendix has three parts. The first part extends our analysis so as to consider the robustness of the results of the repeated game to a relaxation under which past price selections are publicly observed by all firms. The second part derives the expected profit in (13). The third part provides proofs.

7.1 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 retail 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 focus on the incentive constraints that are associated with collusion in advertising. This assumption is not always plausible, however, and we now briefly discuss the robustness of our analysis when this assumption is relaxed.

[45] As Athey and Bagwell (2008) show in their analysis of price collusion, however, when cost shocks are persistent, the advantage of asymmetric equilibria 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. Lee (2009) shows that the potential disadvantage of SPPE may diminish when colluding firms use a contractual device to restrict their incentives to distort private information for their own gain.

In our extended model, each firm observes the realizations of rival firms' past advertising *and* price choices. A firm with cost type θ can then 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 then can be reduced, and new equilibria exist. At the same time, the equilibria that we characterize above - in which firms set their monopoly prices - continue to exist when price histories are public. In the extended model, if firms simply condition their future play on the public history of advertising, then firms again set their monopoly prices.

Formally, in the repeated game with public price histories, we denote a candidate advertising and pricing schedule as (A, ρ), where $\rho(\theta)$ may differ from the monopoly price $p(\theta)$. 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 $\rho(\widehat{\theta})$. To use the Relaxed Program, we let $W(\widehat{\theta}) \equiv \delta[\sup V - \overline{v}(A(\widehat{\theta}), \rho(\widehat{\theta}); A, \rho)]$ and write the interim-stage profit as

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

For simplicity, we assume that A and ρ are continuously differentiable except at a finite number of points where the functions may jump.

The scheme (A, ρ, 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 $\Pi_2(\widehat{\theta}, \theta; A, \rho) = -D(\rho(\widehat{\theta}))M(\widehat{\theta}; A)$ to get

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

Second, a monotonicity condition must hold: $D(\rho(\theta))M(\theta; A)$ must be nonincreasing in θ. This is established by adding two on-schedule incentive constraints:

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

As in Lemma 1, these two necessary conditions are also sufficient for (A, ρ, W) to satisfy on-schedule incentive compatibility.

We now restrict attention to those incentive-compatible schemes (A, ρ, W) for which informed consumers are rational in using the advertising search rule. Given this restriction, we find that $A(\theta)$ must be nonincreasing and $\rho(\theta)$ must be nondecreasing; thus, (A, ρ, 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 $D(\rho(\theta))$ are each nonincreasing.[46] Consider next the potential use of wars. When past prices are public, we cannot immediately use the arguments in Proposition 3 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 $A^*(\theta) \equiv$

[46] Assume to the contrary that $\theta > \widehat{\theta}$ and $A(\theta) > A(\widehat{\theta})$. Given the restriction that informed consumers rationally use the advertising search rule, this assumption implies $M(\theta; A) > M(\widehat{\theta}; A)$ and $\rho(\theta) \leq \rho(\widehat{\theta})$ (i.e., $D(\rho(\theta)) \geq D(\rho(\widehat{\theta}))$). Thus, $A(\theta) > A(\widehat{\theta})$ implies $D(\rho(\theta))M(\theta; A) > D(\rho(\widehat{\theta}))M(\widehat{\theta}; A)$, which contradicts the requirement that $D(\rho(\theta))M(\theta; A)$ is nonincreasing. Hence, $A(\theta)$ must be nonincreasing. Under the restriction that informed consumers rationally use the advertising search rule, if $A(\theta)$ is nonincreasing, then $M(\theta; A)$ is nonincreasing and $\rho(\theta)$ is nondecreasing (i.e., $D(\rho(\theta))$ is nonincreasing).

$A(\theta) + W(\theta)$ would exhibit the necessary nonincreasing property.[47] We can establish that wars are unnecessary in the limiting case where demand is perfectly inelastic; however, the arguments used for this limiting case cannot be directly applied when demand is downward sloping.[48]

We now argue that robust forces in favor of at least partial pooling in advertising are present in the extended model. We develop our argument in three steps. First, we consider the limiting case in which demand is perfectly inelastic and assume that the reservation value r satisfies $r > \overline{\theta}$. As just noted, in this case, wars are unnecessary, and so we focus on stationary SPPE (i.e., schemes (A, ρ, W) in which $W \equiv 0$). Let us now fix any candidate nondecreasing pricing schedule $\rho(\theta)$ satisfying $\rho(\overline{\theta}) > \overline{\theta}$.[49] In the case of perfectly inelastic demand, our monotonicity requirement reduces to the requirement that $M(\theta; A)$ is nonincreasing. Arguing as in the proof of Proposition 5, if F is log-concave, we may establish that expected profit is then maximized by the advertising schedule in which $A(\theta) \equiv 0$, so that $M(\theta; A) \equiv \frac{1}{N}$. With $A(\theta) \equiv 0 \equiv W(\theta)$, there is no potential gain to firms from distorting prices; thus, the optimal nondecreasing pricing schedule entails monopoly pricing with $\rho(\theta) \equiv r$. Thus, for the limiting case in which demand is perfectly inelastic, whether or not rivals' past prices are publicly observed, we can construct an optimal SPPE for patient firms in which firms pool at zero advertising in each period.

Second, returning to our assumption of downward-sloping demand, let us consider any incentive-compatible scheme (A, ρ, W) for which $D(\rho(\theta))$ and $M(\theta; A)$ are nonincreasing, and let us further restrict consideration to pricing functions for which $D(\rho(\theta))\frac{F}{f}(\theta)$ is nondecreasing and $\rho(\overline{\theta}) \geq \overline{\theta}$. Expected profit for incentive-compatible schemes can be represented as

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

If we maximize this expression over the incentive-compatible schemes under consideration, then we may argue as in the proof of Proposition 5 that expected profit is maximized when $M(\theta; A) \equiv \frac{1}{N}$ and thus $A^* \equiv 0$. In a case of special interest, the restriction that $D(\rho(\theta))\frac{F}{f}(\theta)$ is nondecreasing is satisfied if F is log-concave and all types of firms set a constant price, $\overline{\rho} \equiv \rho(\theta) \geq \overline{\theta}$. Firms

[47] 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 $\rho(\theta) = p(\theta_c)$ for types at or above θ_c while $\rho(\theta) = p(\underline{\theta})$ for types below θ_c. Even though market share is higher for lower types, a firm with 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.

[48] Suppose that demand is perfectly inelastic and consider a two-step scheme. The two steps are separated by a critical type, θ_c, and we let θ_b represent a type on the bottom step and θ_t represent a type on the top step (i.e., $\theta_b < \theta_c < \theta_t$). Suppose that $A(\theta_b) > A(\theta_t)$ and thus $M(\theta_b; A) > M(\theta_t; A)$. Suppose further that $A + W$ increases across the steps: $A(\theta_t) + W(\theta_t) > A(\theta_b) + W(\theta_b)$. Incentive compatibility is satisfied if type θ_c is indifferent between the two steps. Given that the top step entails a lower value for M and a higher value for $A + W$, this is possible only if the top step entails a higher price: $\rho(\theta_t) > \rho(\theta_b)$. We now create a new scheme, in which $W(\theta_t)$ is lowered to a new value, $W_N(\theta_t)$, at which $A(\theta_b) + W(\theta_b) = A(\theta_t) + W_N(\theta_t) + \varepsilon$, for $\varepsilon > 0$ small. To maintain incentive compatibility, we adjust $\rho(\theta_t)$ downward until type θ_c is again indifferent. The resulting new price, $\rho_N(\theta_t)$, satisfies $\rho_N(\theta_t) > \rho(\theta_b)$. This maneuver maintains profit for all types. We next eliminate wars and define A^* in terms of the new scheme: $A^*(\theta_b) = A(\theta_b) + W(\theta_b)$ and $A^*(\theta_t) = A(\theta_t) + W_N(\theta_t)$. Note that A^* decreases with θ as we move from the bottom step to the top step, just as did A; hence, A^* generates the same market share allocation as did A. We have thus generated a point-wise equivalent no-war scheme. Finally, we note that, if demand were instead downward sloping, then such step-by-step maneuvers would not generate a point-wise equivalent no-war scheme. This is because the appeal of a price change then varies with cost type. For related reasons, Athey et al. (2004) are also unable to eliminate wars when demand is downward sloping.

[49] If $\rho(\overline{\theta}) \leq \overline{\theta}$ for a nondecreasing price schedule, we could raise $\rho(\overline{\theta})$ above $\overline{\theta}$ and adjust all prices for lower types upward so as to maintain incentive compatibility. This maneuver would raise expected profit, and so we may restrict attention to candidate pricing schedules satisfying $\rho(\overline{\theta}) > \overline{\theta}$.

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 Proposition 3 and show that, if F is log-concave, then optimal SPPE for patient firms entails $A^* \equiv 0$ and $W^* \equiv 0$. Of course, in the case where price is exogenously fixed at $\overline{p}$, it is immaterial whether or not price is public.

Third, robust forces remain in favor of at least partial pooling in advertising even for general demand functions.[50] We make this point in a simple way. Consider any scheme (A, ρ, W) in which A entails full sorting over $[\underline{\theta}, \overline{\theta}]$. We construct an alternative scheme $(A^* \equiv 0, \rho^*, W^* \equiv 0)$, where ρ^* is constant and satisfies

$$\int_{\underline{\theta}}^{\overline{\theta}} D(\rho^*)\frac{1}{N}dF(x) = \int_{\underline{\theta}}^{\overline{\theta}} D(\rho(x))[\frac{U}{N} + [1 - F(x)]^{N-1}I]dF(x).$$

We then define a distribution function under $A^* \equiv 0$ and ρ^*:

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

A distribution $G(\theta; A, \rho)$ is analogously defined under A and ρ. Given that $G(\theta; A^*, \rho^*)$ first-order stochastically dominates $G(\theta; A, \rho)$, if F is log-concave, then the alternative scheme generates higher expected information rents than does the original scheme:

$$\int_{\underline{\theta}}^{\overline{\theta}} \frac{F}{f}(\theta)D(\rho^*)\frac{1}{N}dF(\theta) \geq \int_{\underline{\theta}}^{\overline{\theta}} \frac{F}{f}(\theta)D(\rho(\theta))[\frac{U}{N} + [1 - F(x)]^{N-1}I]dF(\theta).$$

If $\rho^* > \overline{\theta}$ and I is sufficiently large such that $r(\rho^*, \overline{\theta})\frac{1}{N} \geq r(\rho(\overline{\theta}), \overline{\theta})\frac{U}{N}$, then the alternative scheme does not cause any reduction in the profit at the top. As we argue above, however, as a general matter, we cannot directly conclude that wars are unnecessary under general demand functions.

7.2 Derivation of Expected 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. The interim-stage profit for $\theta \leq y$ is

$$R(\theta, \theta; A) - A(\theta) = R(y, y; A) - A(y) + \int_{\theta}^{y} D(p(x))M(x; A)dx, \tag{A1}$$

while the interim-stage profit at y is

$$R(y, y; A) - A(y) = R(\overline{\theta}, \overline{\theta}; A) - A(\overline{\theta}) + \int_{y}^{\overline{\theta}} D(p(x))M(x; A)dx. \tag{A2}$$

[50] Our discussion here builds on Athey et al. (2004).

Using (A1) and (A2), we find the interim-stage profit for $\theta \leq y$:

$$\begin{aligned} R(\theta,\theta;A) - A(\theta) &= R(\overline{\theta},\overline{\theta};A) - A(\overline{\theta}) + \int_{\theta}^{y} D(p(x))M(x;A)dx \\ &\quad + \int_{y}^{\overline{\theta}} D(p(x))M(x;A)dx. \end{aligned} \tag{A3}$$

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

$$R(\theta,\theta;A) - A(\theta) = R(\overline{\theta},\overline{\theta};A) - A(\overline{\theta}) + \int_{\theta}^{\overline{\theta}} D(p(x))M(x;A)dx. \tag{A4}$$

Based on the two interim-stage profits, (A3) and (A4), we find the expected profit (13) by integrating by parts and setting $A(\overline{\theta}) = 0$.

7.3 Proofs

Proof of Lemma 2. We prove that any optimal no-war scheme $(A, W \equiv 0)$ has pooling at the top interval $(y, \overline{\theta}]$ where $A(\theta) = 0$. Suppose that a scheme has a sorting interval at the top on $(y, \overline{\theta}]$. Then we can consider an alternative scheme A^* that decomposes the sorting interval $(y, \overline{\theta}]$ into a sorting interval $(y, y^*]$ and a pooling interval $(y^*, \overline{\theta}]$ where $y^* > y$. The expected profit under A^* becomes

$$\begin{aligned} E_\theta[\Pi(\theta,\theta;A^*)] = r(p(\overline{\theta}),\overline{\theta}) &\left[\frac{U}{N} + [1-F(y^*)]^{N-1}\frac{I}{N}\right] \\ &+ \int_{\underline{\theta}}^{y} D(p(\theta))F(\theta)M(\theta;A^*)d\theta \\ &+ \int_{y}^{y^*} D(p(\theta))F(\theta)\left[\frac{U}{N} + [1-F(\theta)]^{N-1}I\right]d\theta \\ &+ \int_{y^*}^{\overline{\theta}} D(p(\theta))F(\theta)\left[\frac{U}{N} + [1-F(y^*)]^{N-1}\frac{I}{N}\right]d\theta. \end{aligned}$$

Note that, if $y^* \to \overline{\theta}$, then this scheme A^* approaches the initial scheme. We show that the optimal choice of y^* is lower than $\overline{\theta}$. The derivative of expected profit with respect to y^* is given by

$$\begin{aligned} &[1-F(y^*)]^{N-1}\frac{(N-1)I}{N}\left[D(p(y^*))F(y^*) - r(p(\overline{\theta}),\overline{\theta})\frac{f(y^*)}{1-F(y^*)}\right] \\ &-\frac{(N-1)I}{N}\int_{y^*}^{\overline{\theta}} D(p(\theta))F(\theta)[1-F(y^*)]^{N-2}f(y^*)d\theta. \end{aligned}$$

Since $f(\overline{\theta}) > 0$ and $p(\overline{\theta}) > \overline{\theta}$, expected profit rises when y^* slightly falls from $\overline{\theta}$. ■

Proof of Proposition 7. Assume that a no-war scheme A entails sorting for $\theta \in (\theta_i, \theta_{i+1}]$ such that

$$\left\{\theta : A'(\theta) < 0\right\} \cap \left\{\theta : D(p(\theta))\tfrac{F}{f}(\theta) \text{ is nondecreasing}\right\} = (\theta_i, \theta_{i+1}].$$

This interval $(\theta_i, \theta_{i+1}]$ cannot be the interval at the top, given that we restrict attention to the no-war scheme that has pooling at zero advertising on an interval at the top. We define an alternative no-war scheme A^* as:

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

The notations $A_+(\theta)$ and $M_+(\theta; A^*)$ represent the associated limit from the right. 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 $A(\theta_i) - A^*(\theta_i)$ below θ_i. In our notation, $M(\theta; A^*)$ for $\theta \in (\theta_i, \theta_{i+1}]$ equals $M_+(\theta_i; A^*)$.

We first prove that, if $\phi(\theta) \equiv D(p(\theta))\frac{F}{f}(\theta)$ is nondecreasing for $\theta \in (\theta_i, \theta_{i+1}]$, then expected profit is weakly higher under A^* than under A. We define the distribution function under A^*:

$$G(\theta_i, \theta_{i+1}; A^*) \equiv \frac{\int_{\theta_i}^{\theta} M(x; A^*) f(x) dx}{\int_{\theta_i}^{\theta_{i+1}} M(x; A^*) f(x) dx} \text{ for } \theta \in (\theta_i, \theta_{i+1}],$$

where $M(x; A^*)$ represents the market share allocated to $x \in (\theta_i, \theta_{i+1}]$ under pooling:

$$M(x; A^*) = \frac{U}{N} + \sum_{j=0}^{N-1} \binom{N-1}{j} \frac{1}{j+1} \left[F(\theta_{i+1}) - F(\theta_i)\right]^j \left[1 - F(\theta_{i+1})\right]^{N-j-1} I. \tag{A5}$$

The distribution $G(\theta_i, \theta_{i+1}; A)$ is analogously defined under A where $M(x; A) = \frac{U}{N} + [1 - F(x)]^{N-1} I$. We next show that $G(\theta_i, \theta_{i+1}; A^*)$ first-order stochastically dominates $G(\theta_i, \theta_{i+1}; A)$.

To this end, we begin by showing that the two distribution functions have the same denominators:

$$\forall \theta_{i+1} \geq \theta_i, \quad \int_{\theta_i}^{\theta_{i+1}} M(x; A^*) f(x) dx = \int_{\theta_i}^{\theta_{i+1}} M(x; A) f(x) dx. \tag{A6}$$

The equality is immediate if $\theta_{i+1} = \theta_i$. For any $\theta_{i+1} > \theta_i$, we claim that

$$\frac{\partial \int_{\theta_i}^{\theta_{i+1}} M(x; A^*) f(x) dx}{\partial \theta_{i+1}} = \frac{\partial \int_{\theta_i}^{\theta_{i+1}} M(x; A) f(x) dx}{\partial \theta_{i+1}}. \tag{A7}$$

In other words, the expected market shares (denominators) are the same in both schemes at $\theta_{i+1} = \theta_i$, and we claim that they then increase at the same rate as θ_{i+1} rises above θ_i. Given that A^* is pooling on $(\theta_i, \theta_{i+1}]$, the LHS of (A7) is

$$\begin{aligned} \frac{\partial \int_{\theta_i}^{\theta_{i+1}} M(x; A^*) f(x) dx}{\partial \theta_{i+1}} &= \left[\frac{U}{N} + [1 - F(\theta_{i+1})]^{N-1} I\right] f(\theta_{i+1}) \\ &= \frac{\partial \int_{\theta_i}^{\theta_{i+1}} \left[\frac{U}{N} + [1 - F(x)]^{N-1} I\right] f(x) dx}{\partial \theta_{i+1}} \\ &= \frac{\partial \int_{\theta_i}^{\theta_{i+1}} M(x; A) f(x) dx}{\partial \theta_{i+1}}. \end{aligned}$$

Using (A5), the first equality is established by a tedious work of induction for $N \geq 2$. The last term is the RHS of (A7). Our claim that (A7) holds is now established. Hence, the denominators of the two distributions are the same. Further, using $M(x; A) = \frac{U}{N} + [1 - F(x)]^{N-1} I$, we can derive

$$\begin{aligned} \int_{\theta_i}^{\theta_{i+1}} M(x; A^*) f(x) dx = \int_{\theta_i}^{\theta_{i+1}} & M(x; A) f(x) dx \\ &= \left[[1 - F(\theta_i)]^N - [1 - F(\theta_{i+1})]^N\right] \frac{I}{N} \\ &+ [F(\theta_{i+1}) - F(\theta_i)] \frac{U}{N}, \end{aligned} \tag{A8}$$

so that we now have an explicit expression for the common value taken by the denominators of the two distributions.

We next differentiate the two distribution functions with respect to θ. Since the denominators are the same and do not change with θ as seen in (A8), we find

$$\frac{\partial}{\partial \theta} \left[G\left(\theta_i, \theta_{i+1}; A\right) - G(\theta_i, \theta_{i+1}; A^*)\right] = \frac{\left[M(\theta; A) - M(\theta; A^*)\right] f(\theta)}{\int_{\theta_i}^{\theta_{i+1}} M(x; A^*) f(x) dx},$$

where $M(\theta; A) = \frac{U}{N} + [1 - F(\theta)]^{N-1} I$ and $M(\theta; A^*)$ is given by (A5). Given the range $(\theta_i, \theta_{i+1}]$, $M(\theta; A^*)$ crosses $M(\theta; A)$ from below. It then follows that $G(\theta_i, \theta_{i+1}; A^*)$ first-order stochastically dominates $G(\theta_i, \theta_{i+1}; A)$. In other words, $G\left(\theta_i, \theta_{i+1}; A\right) > G(\theta_i, \theta_{i+1}; A^*)$ for $\theta \in (\theta_i, \theta_{i+1})$ and $G\left(\theta_i, \theta_{i+1}; A\right) = G(\theta_i, \theta_{i+1}; A^*)$ for $\theta \in \{\theta_i, \theta_{i+1}\}$.

We next compare the expected profits under the two schemes. Suppose that A^* is represented by K subintervals; note that the original scheme A may be represented by less than K subintervals if A involves sorting consecutively over its neighboring interval, $(\theta_{i-1}, \theta_i]$ or $(\theta_{i+1}, \theta_{i+2}]$. The expected profit under A^* is

$$\begin{aligned} E_\theta \left[\Pi(\theta, \theta; A^*)\right] = & \, r(p(\overline{\theta}), \overline{\theta}) M(\overline{\theta}; A^*) \\ & + \sum_{k \neq i, k=1}^{K} \int_{\theta_k}^{\theta_{k+1}} \phi(\theta) M(\theta; A^*) f(\theta) d\theta \\ & + \int_{\theta_i}^{\theta_{i+1}} \phi(\theta) M(\theta; A^*) f(\theta) d\theta. \end{aligned} \tag{A9}$$

Since A^* is designed to preserve the original market shares under A except for $\theta \in (\theta_i, \theta_{i+1}]$

(as proven below), the expected profit under A is the same as the RHS of (A9) except for the last information-rent term. To compare the associated information-rent terms, we evaluate the differential:

$$\begin{aligned}
&\int_{\theta_i}^{\theta_{i+1}} \phi(\theta) dG(\theta_i, \theta_{i+1}; A^*) - \int_{\theta_i}^{\theta_{i+1}} \phi(\theta) dG\left(\theta_i, \theta_{i+1}; A\right) \\
&= \int_{\theta_i}^{\theta_{i+1}} \phi'(\theta) \left[G\left(\theta_i, \theta_{i+1}; A\right) - G(\theta_i, \theta_{i+1}; A^*)\right] d\theta.
\end{aligned}$$

If $\phi'(\theta) \geq 0$ for $\theta \in (\theta_i, \theta_{i+1}]$, then

$$\int_{\theta_i}^{\theta_{i+1}} \phi(\theta) dG(\theta_i, \theta_{i+1}; A^*) \geq \int_{\theta_i}^{\theta_{i+1}} \phi(\theta) dG\left(\theta_i, \theta_{i+1}; A\right).$$

The inequality can be rewritten as

$$\int_{\theta_i}^{\theta_{i+1}} \phi(\theta) M(\theta; A^*) f(\theta) d\theta \geq \int_{\theta_i}^{\theta_{i+1}} \phi(\theta) M(\theta; A) f(\theta) d\theta.$$

Thus, if $\phi'(\theta) \geq 0$ for $\theta \in (\theta_i, \theta_{i+1}]$, then A^* makes a weakly higher expected profit than A. Further, if $\phi'(\theta) \geq 0$ for $\theta \in (\theta_i, \theta_{i+1}]$ and $\phi'(\theta) > 0$ for some $\theta \in (\theta_i, \theta_{i+1})$, then A^* makes a strictly higher expected profit than A; in this case, an optimal no-war scheme cannot entail sorting for $\theta \in (\theta_i, \theta_{i+1}]$.

We next show that A^* preserves the original market shares under A except for $\theta \in (\theta_i, \theta_{i+1}]$. Consider two cases. Suppose first that the original scheme is sorting consecutively over its neighboring interval, $(\theta_{i-1}, \theta_i]$ or $(\theta_{i+1}, \theta_{i+2}]$, so that A is sorting for $\theta \in (\theta_{i-1}, \theta_{i+1}]$ or for $\theta \in (\theta_i, \theta_{i+2}]$. Then, pooling on $(\theta_i, \theta_{i+1}]$ does not affect the (expected) market share for types on the neighboring sorting interval; in particular, for any sorting interval, the market share for θ is $\frac{U}{N} + [1 - F(\theta)]^{N-1} I$. Suppose second that the original scheme is sorting for $\theta \in (\theta_i, \theta_{i+1}]$ and is adjacent to a pooling interval. If A is pooling for $\theta \in (\theta_{i-1}, \theta_i]$ or for $\theta \in (\theta_{i+1}, \theta_{i+2}]$, then it has a jump at θ_i or at θ_{i+1}. If A is pooling on $(\theta_{i-1}, \theta_i]$, then the market share for $\theta \in (\theta_{i-1}, \theta_i]$ is the same under A and under A^*:

$$\frac{U}{N} + \sum_{j=0}^{N-1} \binom{N-1}{j} \frac{1}{j+1} \left[F(\theta_i) - F(\theta_{i-1})\right]^j \left[1 - F(\theta_i)\right]^{N-j-1} I.$$

Likewise, if A is pooling on $(\theta_{i+1}, \theta_{i+2}]$, then the market share for $\theta \in (\theta_{i+1}, \theta_{i+2}]$ is the same under both schemes. If the original market shares for the neighboring intervals are not affected by A^*, then the market shares for the remaining types will not be affected by A^*.

We finally show that our result holds for any optimal SPPE. Assume that an optimal SPPE exists in which the associated advertising schedule entails sorting for $\theta \in (\theta_i, \theta_{i+1}]$ when $\phi'(\theta) \geq 0$ for $\theta \in (\theta_i, \theta_{i+1}]$ and $\phi'(\theta) > 0$ for some $\theta \in (\theta_i, \theta_{i+1})$. We may translate the factorization of this SPPE into a scheme (A, W) that satisfies the constraints of the Relaxed Program. This scheme in turn is point-wise equivalent to a no-war scheme $(A^*, W^* \equiv 0)$, where A^* is sorting for $\theta \in (\theta_i, \theta_{i+1}]$. As argued above, this no-war scheme can be strictly improved upon by an alternative no-war scheme that is pooling for $\theta \in (\theta_i, \theta_{i+1}]$. If firms are sufficiently patient, we can construct an SPPE in which the advertising schedule from this alternative no-war scheme is used in every period along the equilibrium path. This contradicts our initial assumption.■

Proof of Proposition 8. We show that an optimal SPPE has at most two pooling steps when $\phi(\theta) \equiv D(p(\theta))\frac{F}{f}(\theta)$ is strictly quasiconcave with a maximizer $\theta^* \in (\underline{\theta}, \overline{\theta}]$. When $\theta^* = \overline{\theta}$, the proof is immediate from Proposition 6. Hence, we consider the case in which $\underline{\theta} < \theta^* < \overline{\theta}$. We establish the finding in three steps.

Step 1: We establish two findings. First, an optimal SPPE cannot have two separate pooling steps within $[\underline{\theta}, \theta^*]$ where $\phi'(\theta) > 0$. Suppose that an optimal scheme A has two separate pooling intervals, $[\underline{\theta}, y]$ and $(y, z]$, within $[\underline{\theta}, \theta^*]$ such that $\underline{\theta} < y < z \leq \theta^*$. We construct an alternative scheme A^* that replaces the two pooling steps with one pooling step for $\theta \in [\underline{\theta}, z]$. The original market shares under A are affected by A^* for $\theta \in [\underline{\theta}, z]$. Given the affected range $[\underline{\theta}, z]$, define the distribution function under A:

$$G(\underline{\theta}, z; A) \equiv \frac{\int_{\underline{\theta}}^{\theta} M(x; A) f(x) dx}{\int_{\underline{\theta}}^{y} M(x; A) f(x) dx + \int_{y}^{z} M(x; A) f(x) dx} \text{ for } \theta \in [\underline{\theta}, z]. \tag{A10}$$

The distribution function under A^* is

$$G(\underline{\theta}, z; A^*) \equiv \frac{\int_{\underline{\theta}}^{\theta} M(x; A^*) f(x) dx}{\int_{\underline{\theta}}^{z} M(x; A^*) f(x) dx} \text{ for } \theta \in [\underline{\theta}, z]. \tag{A11}$$

The denominators of distributions are the same. Using (A8), the denominator of (A10) is

$$\begin{aligned}
&\left[[1 - F(\underline{\theta})]^N - [1 - F(y)]^N\right] \frac{I}{N} + [F(y) - F(\underline{\theta})] \frac{U}{N} \\
&+ \left[[1 - F(y)]^N - [1 - F(z)]^N\right] \frac{I}{N} + [F(z) - F(y)] \frac{U}{N} \\
&= \left[[1 - F(\underline{\theta})]^N - [1 - F(z)]^N\right] \frac{I}{N} + [F(z) - F(\underline{\theta})] \frac{U}{N}.
\end{aligned}$$

The RHS is the denominator of (A11). In the range $[\underline{\theta}, z]$, $M(\theta; A^*)$ crosses $M(\theta; A)$ from below and thus $G(\underline{\theta}, z; A^*)$ first-order stochastically dominates $G(\underline{\theta}, z; A)$: $G(\underline{\theta}, z; A) > G(\underline{\theta}, z; A^*)$ for $\theta \in (\underline{\theta}, z)$ and $G(\underline{\theta}, z; A) = G(\underline{\theta}, z; A^*)$ for $\theta \in \{\underline{\theta}, z\}$. It then follows that the expected profit is strictly higher under A^* than under A. This contradict the optimality of A. This non-optimality of the original scheme can be readily extended to any other scheme that has multiple pooling steps or includes some sorting interval within $[\underline{\theta}, z]$.

Second, given $\phi'(\theta) < 0$ for $\theta \in (\theta^*, \overline{\theta}]$, an optimal scheme cannot include a separate pooling interval within $(\theta^*, \overline{\theta}]$ other than at the top. Suppose that a scheme A has a separate pooling step $(\theta^*, y]$ other than the pooling step at the top $(z, \overline{\theta}]$. We construct an alternative scheme A^* that replaces pooling with sorting for $\theta \in (\theta^*, y]$. The original market shares under A are affected by A^* for $\theta \in (\theta^*, y]$. Given the affected range $(\theta^*, y]$, the distribution $G(\theta^*, y; A)$ first-order stochastically dominates $G(\theta^*, y; A^*)$. We then compare the affected information-rent terms:

$$\begin{aligned}
&\int_{\theta^*}^{y} \phi(x) dG(\theta^*, y; A^*) - \int_{\theta^*}^{y} \phi(x) dG(\theta^*, y; A) \\
&= \int_{\theta^*}^{y} \phi'(x) \left[G(\theta^*, y; A) - G(\theta^*, y; A^*)\right] dx > 0.
\end{aligned}$$

Thus, the expected profit is strictly higher under A^* than under A. This contradicts the optimality of A.

Step 2: An optimal SPPE cannot have three pooling steps; in particular, it cannot have an intermediate pooling interval (other than pooling at the bottom and at the top). Since an optimal SPPE cannot include two separate pooling steps or any sorting interval within $[\underline{\theta}, \theta^*]$ where $\phi'(\theta) > 0$, and since it cannot have a separate pooling step within $(\theta^*, \overline{\theta}]$ other than at the top, the only possibility for an optimal SPPE candidate A to have three pooling steps is that the scheme has two pooling steps, $[\underline{\theta}, y]$ and $(y, z]$, such that $\underline{\theta} < y < \theta^* < z < \overline{\theta}$. We construct an alternative scheme A^* that has two pooling steps, $[\underline{\theta}, y^*]$ and $(y^*, z]$, such that $\underline{\theta} < \theta^* < y^* < z < \overline{\theta}$. Observe that

$$\begin{aligned} M(\theta; A^*) &< M(\theta; A) \text{ for } \theta \in [\underline{\theta}, y] \text{ and } \theta \in (y^*, z] \\ M(\theta; A^*) &> M(\theta; A) \text{ for } \theta \in (y, y^*]. \end{aligned} \tag{A12}$$

For the affected range $[\underline{\theta}, z]$, define the distributions, $G(\underline{\theta}, z; A^*)$ and $G(\underline{\theta}, z; A)$. Since the two functions have the same fixed denominators, we find

$$\frac{\partial}{\partial \theta} \left[G(\underline{\theta}, z; A) - G(\underline{\theta}, z; A^*) \right] = \frac{\left[M(\theta; A) - M(\theta; A^*) \right] f(\theta)}{\int_{\underline{\theta}}^{z} M(x; A^*) f(x) dx}.$$

Given the inequalities in (A12), the slope of $G(\theta, z; A^*)$ is flatter (steeper) than that of $G(\underline{\theta}, z; A)$ at the bottom $[\underline{\theta}, y]$ and at the top $(y^*, z]$ (at the intermediate range $(y, y^*]$). Thus, in the range $[\underline{\theta}, z]$, $G(\underline{\theta}, z; A^*)$ crosses $G(\underline{\theta}, z; A)$ from below other than at the two endpoints, $\underline{\theta}$ and z, where $G(\underline{\theta}, z; A^*) = G(\underline{\theta}, z; A)$. The choice of the point y^* is made to satisfy $\theta^* < y^* < z$. If $y^* \leq \theta^*$, then $G(\underline{\theta}, z; A^*)$ crosses $G(\underline{\theta}, z; A)$ at type $\theta < \theta^*$, since the slope of $G(\underline{\theta}, z; A^*)$ is flatter than that of $G(\underline{\theta}, z; A)$ in the range $(y^*, z]$. If $y^* = z$, then $G(\underline{\theta}, z; A^*)$ crosses $G(\underline{\theta}, z; A)$ at the endpoint $z > \theta^*$; if $y^* = z$, then $G(\underline{\theta}, z; A^*)$ first-order stochastically dominates $G(\underline{\theta}, z; A)$: $G(\underline{\theta}, z; A^*) < G(\underline{\theta}, z; A)$ for all $\theta \in (\underline{\theta}, z)$. It is thus possible to adjust the level of y^* such that $G(\underline{\theta}, z; A^*)$ crosses $G(\underline{\theta}, z; A)$ from below at θ^*: $G(\underline{\theta}, z; A^*) < G(\underline{\theta}, z; A)$ for $\theta \in (\underline{\theta}, \theta^*)$ and $G(\underline{\theta}, z; A^*) > G(\underline{\theta}, z; A)$ for $\theta \in (\theta^*, z)$. Given the choice of y^*, we can compare the two affected information-rent terms:

$$\begin{aligned} &\int_{\underline{\theta}}^{z} \phi(x) dG(\underline{\theta}, z; A^*) - \int_{\underline{\theta}}^{z} \phi(x) dG(\underline{\theta}, z; A) \\ &= \int_{\underline{\theta}}^{\theta^*} \phi'(x) \left[G(\underline{\theta}, z; A) - G(\underline{\theta}, z; A^*) \right] dx \\ &+ \int_{\theta^*}^{z} \phi'(x) \left[G(\underline{\theta}, z; A) - G(\underline{\theta}, z; A^*) \right] dx > 0. \end{aligned}$$

The expected profit is strictly higher under A^* than under A. This is a contradiction for A to be optimal. Further, note that, other than the pooling at the top, A^* now includes a separate pooling step $(y^*, z]$ within $(\theta^*, \overline{\theta}]$ where $\phi'(\theta) < 0$. We can then construct another alternative scheme that replaces pooling with sorting for $\theta \in (y^*, z]$ and strictly improves upon the scheme A^*. Hence, if $\phi(\theta)$ is strictly quasiconcave with a maximizer $\theta^* \in (\underline{\theta}, \overline{\theta}]$, then an optimal SPPE cannot have an intermediate pooling interval.

Step 3: If an intermediate sorting interval, $(y, z]$, is ever used, then it is restricted to a subset of the range in which $\phi'(\theta) < 0$: $\theta^* < y < z < \overline{\theta}$. If a scheme A has three intervals, $[\underline{\theta}, y]$, $(y, z]$ and $(z, \overline{\theta}]$, such that $y \le \theta^*$, then we can construct an alternative scheme A^* that has three intervals, $[\underline{\theta}, y^*]$, $(y^*, z]$ and $(z, \overline{\theta}]$, such that $y^* > \theta^*$. The alternative scheme lengthens the first pooling step *beyond* the range $[\underline{\theta}, \theta^*]$. The original market shares under A are affected by A^* for $\theta \in [\underline{\theta}, y^*]$. Given the affected types $\theta \in [\underline{\theta}, y^*]$, define the distributions, $G(\underline{\theta}, y^*; A^*)$ and $G(\underline{\theta}, y^*; A)$. Since A^* has one pooling step, $M(\theta; A^*)$ crosses $M(\theta; A)$ from below, and $G(\underline{\theta}, y^*; A^*)$ first-order stochastically dominates $G(\underline{\theta}, y^*; A)$. We then compare the affected information-rent terms:

$$\begin{aligned}
&\int_{\underline{\theta}}^{y^*} \phi(\theta) dG(\underline{\theta}, y^*; A^*) - \int_{\underline{\theta}}^{y^*} \phi(\theta) dG\left(\underline{\theta}, y^*; A\right) \\
&= \int_{\underline{\theta}}^{\theta^*} \phi'(\theta) \left[G\left(\underline{\theta}, y^*; A\right) - G(\underline{\theta}, y^*; A^*)\right] d\theta \\
&+ \int_{\theta^*}^{y^*} \phi'(\theta) \left[G\left(\underline{\theta}, y^*; A\right) - G(\underline{\theta}, y^*; A^*)\right] d\theta.
\end{aligned}$$

The first term on the RHS is positive and the second term is negative. We can choose y^* slightly above θ^* so that the RHS becomes strictly positive. This finding is confirmed by differentiation of the RHS with respect to y^*:

$$\int_{\underline{\theta}}^{\theta^*} \phi'(\theta) \frac{\partial \Delta(\theta)}{\partial y^*} d\theta + \int_{\theta^*}^{y^*} \phi'(\theta) \frac{\partial \Delta(\theta)}{\partial y^*} d\theta + \phi'(y^*) \Delta(y^*),$$

where $\Delta(\theta) \equiv G\left(\underline{\theta}, y^*; A\right) - G(\underline{\theta}, y^*; A^*)$. Note that $\Delta(y^*) = 0$, and so the third term is zero. If y^* approaches θ^* from the right, then the first term remains positive given $\frac{\partial \Delta(\theta)}{\partial y^*} > 0$, while the second term approaches zero. ■

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

Athey, S. and K. Bagwell (2008), "Collusion with Persistent Cost Shocks," *Econometrica*, 76, 493-540.

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. M. Lee (2009), "Advertising Competition in Retail Markets," mimeo.

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

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.

California Dental Association v. Federal Trade Commission (1999), No. 97-1625, May 24.

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.

Fudenberg, D., Levine, D. J., and E. Maskin (1994), "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997-1039.

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.

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.

LeBlanc, G. (1998), Informative Advertising Competition," *Journal of Industrial Economics*, 46, 63-77.

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

Lee, G. M. (2009), "Optimal Collusion with Internal Contracting," forthcoming, *Games and Economic Behavior.*

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.

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

Peters, M. (1984), "Restrictions on Price Advertising," *Journal of Political Economy*, 92, 472-85.

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.

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.