Merger Policy with Merger Choice*

Volker Nocke
University of Mannheim, CESifo and CEPR

Michael D. Whinston
Northwestern University and NBER

March 19, 2011

Abstract

We analyze the optimal policy of an antitrust authority towards horizontal mergers when merger proposals are endogenous and firms choose which of several mutually exclusive mergers to propose. The optimal policy of an antitrust authority that seeks to maximize expected consumer surplus involves discriminating between mergers based on a naive computation of the post-merger Herfindahl index (over and above the apparent effect of the proposed merger on consumer surplus). We show that the antitrust authority optimally imposes a tougher standard on those mergers that raise the index by more.

1 Introduction

The evaluation of proposed horizontal mergers involves a basic trade-off: mergers may increase market power, but may also create efficiencies. Whether a given merger should be approved depends, as first emphasized by Williamson (1968), on a balancing of these two effects.

In most of the literature discussing horizontal merger evaluation, the assumption is that a merger should be approved if and only if it improves welfare, whether that be aggregate surplus or just consumer surplus, as is in practice the standard adopted by most antitrust authorities [see, e.g., Farrell and Shapiro (1990), McAfee and Williams (1992)]. This paper contributes to a small literature that formally derives optimal merger approval rules. This literature started with Besanko and Spulber (1993), who discussed the optimal rule for an antitrust authority who cannot directly observe efficiencies but who recognizes that firms know this information and decide whether to propose a merger based on this knowledge. Other recent papers in this literature include Armstrong and Vickers (2010), Nocke and Whinston (2010), and Ottaviani and Wickelgren (2009).

In this paper, we focus on a setting in which one "pivotal" firm may merge with one of a number of other firms who have differing initial marginal cost levels. These mergers are mutually exclusive,

*We thank members of the Toulouse Network for Information Technology, Nuffield College's econonomic theory lunch, and various seminar audiences for their comments. Nocke gratefully acknowledges financial support from the UK's Economic and Social Research Council, as well as the hospitality of Northwestern University's Center for the Study of Industrial Organization. Whinston thanks the National Science Foundation, the Toulouse Network for Information Technology, and the Leverhulme Trust for financial support, as well as Nuffield College and the Oxford University Department of Economics for their hospitality.

and each may result in a different, randomly drawn post-merger marginal cost due to merger-related synergies. The merger that is proposed is the result of a bargaining process among the firms. The antitrust authority observes the characteristics of the merger that is proposed, but neither the feasibility nor the characteristics of any mergers that are not proposed.

We focus in the main part of the paper on an antitrust authority who wishes to maximize expected consumer surplus. Our main result characterizes the form of the antitrust authority's optimal policy, which we show should impose a tougher standard on mergers involving larger merger partners (in terms of their pre-merger market share). Specifically, the minimal acceptable improvement in consumer surplus is strictly positive for all but the smallest merger partner, and is larger the greater is the merger partner's pre-merger share. Since in this baseline model a greater pre-merger share for the merger partner is equivalent to a larger naively-computed post-merger Herfindahl index, another way to say this is that mergers that result in a larger naively-computed post-merger Herfindahl index must generate larger improvements in consumer surplus to be approved.[1]

The closest papers to ours are Lyons (2003) and Armstrong and Vickers (2010). Lyons is the first to identify the issue that arises when firms may choose which merger to propose. Armstrong and Vickers (2010) provide an elegant characterization of the optimal policy when mergers (or, more generally, projects that may be proposed by an agent) are ex ante identical in terms of their distributions of possible outcomes. Our paper differs from Armstrong and Vickers (2010) primarily in its focus on the optimal treatment of mergers that differ in this ex ante sense. Moreover, a key issue in our paper – the bargaining process among firms – is absent in Armstrong and Vickers as they consider the case of a single agent.[2]

The paper is also related to Nocke and Whinston (2010). That paper establishes conditions under which the optimal dynamic policy for an antitrust authority who wants to maximize discounted expected consumer surplus is a completely myopic policy, in which a merger is approved if and only if it does not lower consumer surplus at the time it is proposed. A key assumption for that result is that potential mergers are "disjoint," in the sense that the set of firms involved in different possible mergers do not overlap. The present paper explores, in a static setting, the implications of relaxing that disjointness assumption, focusing on the polar opposite case in which all potential mergers are mutually exclusive.

The paper proceeds as follows. We describe the baseline model in Section 2. In Section 3, we derive our main result: the antitrust authority optimally imposes a tougher standard, in terms of the minimum increase in consumer surplus required for approval, the "larger" is the proposed merger. In Section 4, we show that the optimal policy may not have a cutoff structure and provide a sufficient condition under which it does. Assuming it does, we derive some comparative statics results.

In Section 5, we explore several extensions of the baseline model. First, we show that our main result for the baseline model, where we assume that the bargaining between firms proceeds as in the

[1] The naively-computed Herfindahl index is computed assuming that the merged firm's post-merger share is the sum of the merger partners' pre-merger shares and that the shares of outsiders do not change. The change in the Herfindahl index due to the merger between the pivotal firm 0 and firm k, computed in this naive way, equals $2s_0s_k$ – twice the product of the merging firms' pre-merger market shares – and so is larger the greater is s_k. It is interesting to note that, in the U.S. merger guidelines, this naively-computed post-merger Herfindahl index plays a central (although different) role in screening mergers.

[2] From a theory point of view, our paper contributes to the literature on (constrained) delegated agency without transfers, which was initiated by Holmstrom (1984). Recent contributions include Alonso and Matouschek (2008), Armstrong and Vickers (2010), and Che, Dessein, and Kartik (2010). A key difference between Che, Dessein, and Kartik (2010) and our paper is that they assume that the principal (antitrust authority) can condition its policy only on the identity of the proposed project (merger) but not on its characteristics (post-merger costs).

Segal (1999) offer game, extends to other bargaining models. Second, we relax the assumption that firm 0 is a party to any merger and that any merger involves two firms. We show that in this more general setting, the key criterion according to which the antitrust authority should optimally discriminate between alternative mergers is the naively-computed post-merger Herfindahl index. Third, we show that our main result continues to hold if the antitrust authority seeks to maximize aggregate surplus, or any convex combination between consumer surplus and aggregate surplus. Fourth, adopting an aggregative game approach [e.g., Dubey, Haimanko and Zapechelnyuk (2006)], we extend the model to the case of price competition with differentiated products (CES and multinomial logit demand structures). Fifth, we extend the baseline model to allow for fixed cost savings.

We conclude in Section 6.

2 The Model

We consider a homogeneous goods industry in which firms compete in quantities (Cournot competition). Let $\mathcal{N} = \{0, 1, 2, ..., N\}$ denote the (initial) set of firms. All firms have constant returns to scale; firm i's marginal cost is denoted c_i. Inverse demand is given by $P(Q)$. We impose standard assumptions on demand:

Assumption 1. *For all Q such that $P(Q) > 0$, we have:*

(i) $P'(Q) < 0$*;*

(ii) $P'(Q) + QP''(Q) < 0$;

(iii) $\lim_{Q\to\infty} P(Q) = 0$.

It is well known that under these conditions there exists a unique Nash equilibrium in quantities. Moreover, this equilibrium is "stable" [each firm i's best-response function $b_i(Q_{-i}) \equiv \arg\max_{q_i} [P(Q_{-i}+q_i) - c_i]q_i$ satisfies $b_i'(Q_{-i}) \in (-1, 0)$, where $Q_{-i} \equiv \sum_{j\neq i} q_j]$ so that comparative statics are "well behaved" (if a subset of firms jointly produce less [more] because of a change in their incentives to produce output, then equilibrium industry output will fall [rise]). The vector of output levels in the pre-merger equilibrium is given by $q^0 \equiv (q_0^0, q_1^0, ..., q_N^0)$, where q_i^0 is firm i's quantity. For simplicity, we assume that pre-merger marginal costs are such that all firms in $\mathcal{N}$ are "active" in the pre-merger equilibrium, i.e., $q_i^0 > 0$ for all i. Hence, each firm i's output ($i = 0, 1, ..., N$) satisfies the first-order condition

$$P(Q^0) + P'(Q^0)q_i^0 = c_i. \tag{1}$$

Aggregate output, price, consumer surplus, and firm i's profit in the pre-merger equilibrium are denoted $Q^0 \equiv \sum_i q_i^0$, $P^0 \equiv P(Q^0)$, CS^0, and $\pi_i^0 \equiv [P(Q^0) - c_i]q_i^0$, respectively. Firm i's market share is $s_i^0 \equiv q_i^0/Q^0$.

Suppose that there is a set of K potential mergers, each between firm 0 (the "acquirer") and a single merger partner (a "target") $k \in \mathcal{K} \subseteq \mathcal{N}$. There is a random variable $\phi_k \in \{0, 1\}$ that determines whether the merger between firm 0 and firm k is feasible ($\phi_k = 1$) or not ($\phi_k = 0$). We let $\theta_k \equiv \Pr(\phi_k = 1) > 0$ denote the probability that the merger is feasible. A feasible merger is described by $M_k = (k, \overline{c}_k)$, where k is the identity of the target and $\overline{c}_k$ the (realized) post-merger marginal cost, which is drawn from distribution function G_k with support $[l, h_k]$ and no mass points. The random draws of ϕ_k and $\overline{c}_k$ are independent across mergers.

If merger M_k is implemented, the vector of outputs in the resulting post-merger equilibrium is denoted $q(M_k) \equiv (q_1(M_k), ..., q_N(M_k))$, where $q_k(M_k)$ is the output of the merged firm, aggregate output is $Q(M_k) \equiv \sum_i q_i(M_k)$, and firm i's market share is $s_i(M_k) \equiv q_i(M_k)/Q(M_k)$. We assume that all nonmerging firms remain active after any merger, so individual outputs satisfy the first-order condition

$$P(Q(M_k)) + P'(Q(M_k))q_i(M_k) = c_i \tag{2}$$

for the nonmerging firms $i \neq 0, k$ and

$$P(Q(M_k)) + P'(Q(M_k))q_k(M_k) = \overline{c}_k \tag{3}$$

for the merged firm. The post-merger profit of nonmerging firm i is given by $\pi_i(M_k) \equiv [P(Q(M_k)) - c_i]\, q_i(M_k)$, and the merged firm's profit by $\pi_k(M_k) \equiv [P(Q(M_k)) - \overline{c}_k]\, q_k(M_k)$. The induced change in consumer surplus is

$$\Delta CS(M_k) \equiv \left\{ \int_0^{Q(M_k)} P(s)ds - P(Q(M_k))Q(M_k) \right\} - CS^0.$$

We will say that a merger M_k is *CS-neutral* if $\Delta CS(M_k) = 0$, CS-increasing if $\Delta CS(M_k) > 0$, and CS-decreasing if $\Delta CS(M_k) < 0$. A merger is CS-nondecreasing (resp., non-increasing) if it is not CS-decreasing (resp., CS-increasing). If no merger is implemented, the status quo (or "null merger") M_0 obtains, resulting in outcome $q(M_0) \equiv q^0$, $s_i(M_0) \equiv q_i^0/Q^0$, and $\Delta CS(M_0) = 0$. The realized set of feasible mergers is denoted $\mathfrak{F} \equiv \{M_k : \phi_k = 1\} \cup M_0$.

We assume that if merger M_k, $k \in \mathfrak{F}$, is proposed, the antitrust authority can observe all aspects of that merger. We also assume that the antitrust authority can commit ex ante to a merger-specific approval policy by specifying an approval (or "acceptance") set $\mathcal{A} \equiv \{M_k : \overline{c}_k \in \mathcal{A}_k\} \cup M_0$, where $\mathcal{A}_k \subseteq [l, h_k]$ for $k \in \mathcal{K}$ are the post-merger marginal cost levels that would lead to approval of a merger with target k. Because of our assumption of full support and no mass points, we can without loss of generality restrict attention to the case where each $\mathcal{A}_k$ is a (finite or infinite) union of closed intervals, i.e., $\mathcal{A}_k \equiv \cup_{r=1}^{R} [l_k^r, h_k^r]$, where $l \leq l_k^r < h_k^r \leq h_k$ (R can be infinite). Note that the status quo M_0 is always "approved."

Some remarks are in order concerning the policies that we consider: First, we confine attention to deterministic policies. One justification is that it may be hard for the antitrust authority to commit to a random rule. Second, we do not pursue a mechanism design approach. Motivated by the constraints that antitrust authorities face in the real world, we assume that the antitrust authority cannot ask firms for information on mergers that are not proposed. Moreover, we assume that only one of the mutually exclusive mergers can be proposed to, and evaluated by, the antitrust authority.[3]

Given a realized set of feasible mergers $\mathfrak{F}$ and the antitrust authority's approval set $\mathcal{A}$, the set of feasible mergers that would be approved if proposed is given by $\mathfrak{F} \cap \mathcal{A}$. A bargaining process among the firms determines which feasible merger is actually proposed. Note that this bargaining problem involves externalities as firms' payoffs depend on the identity of the target. There are various ways in which one could model this situation. For now, we suppose the bargaining process takes the form of an

[3] In some special cases, the antitrust authority could not do better if we relaxed the assumption that at most one merger can be proposed and evaluated. In particular, suppose firm 0 has private information about the set of feasible mergers (and the efficiencies of these mergers). Further, suppose that the antitrust authority can verify claimed efficiencies only once a merger has been implemented. Finally, suppose there is an independent legal system that would punish any firm for lying to the antitrust authority and that such punishment would outweigh any gain from merging. In that case, there is no welfare loss in our model from restricting firms to propose at most one merger to the antitrust authority.

offer game, as in Segal (1999), where the acquirer (firm 0) – Segal's principal – makes public take-it-or-leave-it offers. In Segal (1999), the principal's offers consist of a profile of "trades" $x = (x_1, ..., x_K)$, with x_k the trade with agent k. Here, $x_k \in \{0, 1\}$, where $x_k = 1$ if the acquirer proposes a merger with firm k. Hence, here Segal's offer game simply amounts to firm 0 being able to make a take-it-or-leave-it offer of an acquisition price t_k to a single firm k of its choosing, where k is such that $M_k \in (\mathfrak{F} \cap \mathcal{A})$. If the offer is accepted by firm k, then merger M_k is proposed to the antitrust authority, who will approve it since $M_k \in (\mathfrak{F} \cap \mathcal{A})$, and firm 0 acquires the target in return for the payment t_k. If the offer is rejected, or if no offer is made, then no merger is proposed and no payments are made. (In Section 5.1 we will discuss other bargaining processes.)

For $k \in \mathcal{K}$, let

$$\Delta\Pi(M_k) \equiv \pi_k(M_k) - \left[\pi_0^0 + \pi_k^0\right],$$

denote the change in the bilateral profit of the merging parties, firms 0 and k, induced by merger $M_k \in (\mathfrak{F} \cap \mathcal{A})$. By choosing the payment t_k that makes firm k just indifferent between accepting and not, firm 0 can extract the entire bilateral profit gain $\Delta\Pi(M_k)$. Given the realized set of feasible and acceptable mergers, $\mathfrak{F} \cap \mathcal{A}$, the proposed merger in the equilibrium of the offer game is therefore given by $M^*(\mathfrak{F}, \mathcal{A})$, where

$$M^*(\mathfrak{F}, \mathcal{A}) \equiv \begin{cases} \arg\max_{M_k \in (\mathfrak{F} \cap \mathcal{A})} \Delta\Pi(M_k) & \text{if } \max_{M_k \in (\mathfrak{F} \cap \mathcal{A})} \Delta\Pi(M_k) > 0 \\ M_0 & \text{otherwise.} \end{cases}$$

That is, the proposed merger M_k is the one that maximizes the induced change in the bilateral profit of firms 0 and k, provided that change is positive; otherwise, no merger is proposed.

In line with legal standards in the U.S., the EU, and many other jurisdictions, we assume that the antitrust authority acts in the consumers' interests. That is, the antitrust authority selects the approval set $\mathcal{A}$ that maximizes expected consumer surplus given that firms' proposal rule is $M^*(\cdot)$:

$$\max_{\mathcal{A}} E_{\mathfrak{F}}\left[\Delta CS\left(M^*(\mathfrak{F}, \mathcal{A})\right)\right],$$

where the expectation is taken with respect to the set of feasible mergers, $\mathfrak{F}$. (We discuss alternative welfare standards in Section 5.4.)

We are interested in studying how the optimal approval set depends on the pre-merger characteristics of the alternative mergers. For this reason, we assume that the potential targets differ in their pre-merger marginal costs. Without loss of generality, let $\mathcal{K} \equiv \{1, ..., K\}$ and re-label firms 1 through K in decreasing order of their pre-merger marginal costs: $c_1 > c_2 > ... > c_K$. Thus, in the pre-merger equilibrium, firm $k \in \mathcal{K}$ produces more than firm $j \in \mathcal{K}$, and has a larger market share, if $k > j$. We will say that merger M_k is *larger* than merger M_j if $k > j$ as the combined pre-merger market share of firms 0 and k is larger than that of firms 0 and j. As noted earlier, in this setting the change in the naively-computed Herfindahl index from a merger between firms 0 and k is $2s_0^0 s_k^0$.[4] Thus, a larger

[4] Specifically, the change in the naively-computed Herfindahl index induced by merger M_k, denoted $\Delta H^{naive}(M_k)$, is given by

$$\begin{aligned} \Delta H^{naive}(M_k) &\equiv \left(\sum_{i \neq 0,k} \left(s_i^0\right)^2 + (s_0^0 + s_k^0)^2\right) - \sum_{i=0}^{N} \left(s_i^0\right)^2 \\ &= (s_0^0 + s_k^0)^2 - (s_0^0)^2 - \left(s_k^0\right)^2 \\ &= 2s_0^0 s_k^0. \end{aligned}$$

merger causes a larger change in this naively-computed index.

3 Optimal Merger Policy

We now investigate the form of the antitrust authority's optimal policy when the bargaining process among firms takes the form of the offer game. Given a realized set of feasible mergers $\mathfrak{F}$ and an approval set $\mathcal{A}$, this bargaining process results in the merger $M^*(\mathfrak{F},\mathcal{A})$, as discussed in the previous section. We begin with some preliminary observations before turning to our main result.

3.1 Preliminaries

As firms produce a homogeneous good, a merger M_k raises (resp. reduces) consumer surplus if and only if it raises (resp. reduces) aggregate output Q. The following lemma summarizes some useful properties of a *CS-neutral* merger M_k, i.e., a merger that leaves consumer surplus unchanged, $\Delta CS(M_k) = 0$.

Lemma 1. *Suppose merger M_k is CS-neutral. Then*

(i) *the merger causes no changes in the output of any nonmerging firm $i \notin \{0, k\}$ nor in the joint output of the merging firms 0 and k;*

(ii) *the merged firm's margin at the pre- and post-merger price $P(Q^0)$ equals the sum of the merging firms' pre-merger margins:*

$$P(Q^0) - \overline{c}_k = \left[P(Q^0) - c_0\right] + \left[P(Q^0) - c_k\right]; \tag{4}$$

(iii) *the merger is profitable for the merging firms: $\Delta\Pi(M_k) > 0$;*

(iv) *the merger increases aggregate profit: $\sum_{i\in\mathcal{N}\setminus\{0\}} \pi_i(M_k) > \sum_{i\in\mathcal{N}} \pi_i^0$.*

Proof. Part (i) follows from stability of equilibrium; part (ii) from the merged firm's first-order condition for profit maximization (3) and from summing the merger partners' pre-merger first-order conditions (1); part (iii) is an implication of parts (i) and (ii) [see Nocke and Whinston (2010) for details]. As for part (iv), note that the merger raises the billateral (i.e., joint) profit of the merging firms 0 and k by part (3) and it leaves the profit of any nonmerging firm unchanged (as neither price nor their output changes). □

Rewriting equation (4), merger M_k is CS-neutral if the post-merger marginal cost $\overline{c}_k$ satisfies

$$\overline{c}_k = \widehat{c}(Q^0) \equiv c_k - \left[P(Q^0) - c_0\right]. \tag{5}$$

An implication of condition (5), emphasized by Farrell and Shapiro (1990), is that a CS-neutral merger must involve a reduction in marginal cost below the marginal cost level of the more efficient merger partner: i.e., M_k can be CS-neutral only if $\overline{c}_k < \min\{c_0, c_k\}$.

As the following standard lemma (proof omitted) shows, reducing the merged firm's marginal cost $\overline{c}_k$ not only increases consumer surplus but also the profit of the merged firm:

Lemma 2. *Conditional on merger M_k being implemented, a reduction in the post-merger marginal cost $\overline{c}_k$ causes aggregate output, consumer surplus, and the merged firm's profit to increase.*

Thus, conditional on merger M_k being implemented, both $\Delta CS(M_k)$ and $\Delta\Pi(M_k)$ – the changes in consumer surplus and bilateral profit of the merging firms – increase when the post-merger marginal cost declines. Combined with (5), this also implies that merger M_k is CS-increasing if $\overline{c}_k < \widehat{c}(Q^0)$ and CS-decreasing if $\overline{c}_k > \widehat{c}(Q^0)$.

To make the antitrust authority's problem interesting, and avoid certain degenerate cases, we will henceforth assume the following:

Assumption 2. *For all $k \in \mathcal{K}$, the support of the post-merger cost distribution includes both CS-increasing and CS-nonincreasing mergers: i.e., $\Delta CS(k, h_k) \leq 0 < \Delta CS(k, l)$.*

The following lemma gives a key result that indicates that there is a systematic bias in the proposal incentives of firms in favor of larger mergers, relative to the interests of consumers:

Lemma 3. *Suppose two mergers, M_j and M_k, with $k > j \geq 1$, induce the same non-negative change in consumer surplus, $\Delta CS(M_j) = \Delta CS(M_k) \geq 0$. Then the larger merger M_k induces a greater increase in the bilateral profit of the merger partners: i.e., $\Delta\Pi(M_k) > \Delta\Pi(M_j) > 0$.*

Proof. Note first that, conditional on aggregate output being Q, firms' first-order conditions (1), (2), and (3) imply that we can write the profit of a firm with marginal cost c_i as

$$\pi_i = -P'(Q)[r(Q; c_i)]^2, \tag{6}$$

where

$$r(Q; c) \equiv \{q : P(Q) - c + P'(Q)q = 0\} = -\frac{P(Q) - c}{P'(Q)}$$

is the "cumulative best reply" of a firm with marginal cost c when aggregate output is Q. Observe also that this function is decreasing in both Q and c. Next, note that adding all firms' first-order conditions implies that the equilibrium aggregate output depends only on the sum of active firms' marginal costs. Thus, since both mergers induce the same aggregate output, $\overline{Q} \equiv Q(M_k) = Q(M_j)$, both mergers involve the same cost saving $\gamma \equiv c_k - \overline{c}_k = c_j - \overline{c}_j$. In fact, any merger between firm 0 and a firm with pre-merger marginal cost c that results in a post-merger marginal cost of $c - \gamma$ expands output from Q^0 to $\overline{Q}$.

The difference in the merged firms' profits can thus be written as

$$\begin{aligned}
\pi_k(M_k) - \pi_j(M_j) &= -P'(\overline{Q})\left\{\left[r(\overline{Q}; \overline{c}_k)\right]^2 - \left[r(\overline{Q}; \overline{c}_j)\right]^2\right\} \\
&= -P'(\overline{Q})\left\{\left[r(\overline{Q}; c_k - \gamma)\right]^2 - \left[r(\overline{Q}; c_j - \gamma)\right]^2\right\} \\
&= -P'(\overline{Q})\int_{c_j}^{c_k} 2r(\overline{Q}; c - \gamma)\frac{\partial r(\overline{Q}; c - \gamma)}{\partial c}dc \\
&= 2\int_{c_k}^{c_j} r(\overline{Q}; c - \gamma)dc.
\end{aligned} \tag{7}$$

Similarly, the difference in the firms' pre-merger profits is given by

$$\pi_k^0 - \pi_j^0 = 2\int_{c_k}^{c_j} r(Q^0; c)dc. \tag{8}$$

Since the merger of firm 0 and a firm with pre-merger marginal cost c that results in a post-merger marginal cost of $c - \gamma$ weakly expands output from Q^0 to $\overline{Q} \geq Q^0$, it must weakly reduce nonmerging firms' outputs and weakly expand the output of the merging firms. Thus,

$$r(\overline{Q}; c - \gamma) \geq r(Q^0, c_0) + r(Q^0, c) > r(Q^0, c).$$

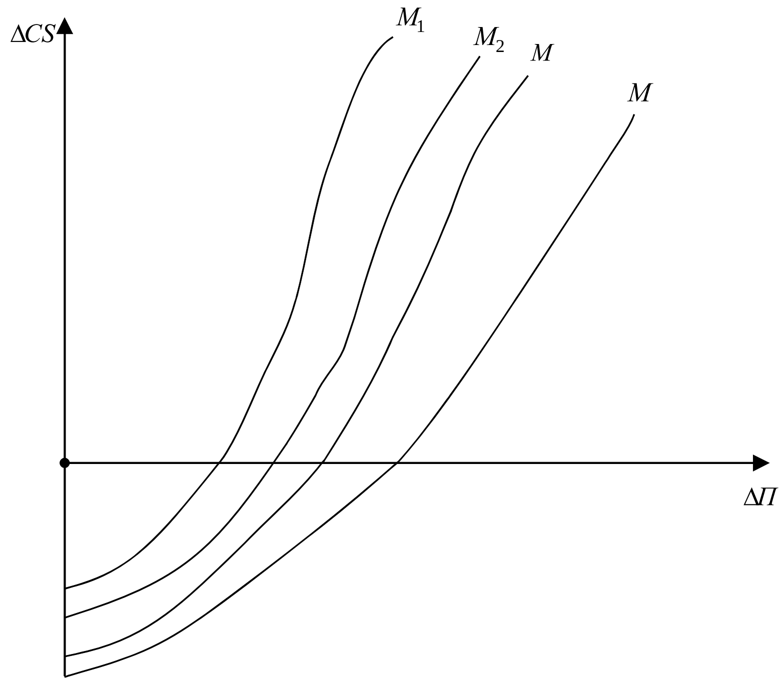

Figure 1: The curves depict the relationship between the change in consumer surplus and the change in bilateral profit of the various mergers, where each point on a curve corresponds to a different realization of post-merger marginal cost for that merger.

By (7) and (8), this implies that $\pi_k(M_k) - \pi_j(M_j) > \pi_k^0 - \pi_j^0$, which can be rewritten as $\Delta\Pi(M_k) > \Delta\Pi(M_j)$. $\square$

Lemmas 1 to 3 imply that the possible mergers can be represented as shown in Figure 1 (where there are four possible mergers; i.e., $K = 4$). In the figure, the change in the merging firms' bilateral profit, $\Delta\Pi$, is measured on the horizontal axis and the change in consumer surplus, ΔCS, is measured on the vertical axis. The CS-increasing mergers therefore are those lying above the horizontal axis. The bilateral profit and consumer surplus changes induced by a merger between firms 0 and $k \geq 1$, $(\Delta\Pi(M_k), \Delta CS(M_k))$, fall somewhere on the curve labeled "M_k." (The figure shows only the parts of these curves for which the bilateral profit change $\Delta\Pi$ is nonnegative.) Since by Lemma 1 a CS-neutral merger is profitable for the merger partners, each curve crosses the horizontal axis to the right of the vertical axis. By Lemma 2, the curve for each merger M_k, $k \geq 1$, is upward sloping. By Lemma 3, on and above the horizontal axis the curves for larger mergers lie everywhere to the right of those for smaller mergers.

A useful corollary of Lemmas 2 and 3, which can easily be seen in Figure 1, is the following:

Corollary 1. *If two CS-nondecreasing mergers M_j and M_k with $k > j \geq 1$ have $\Delta\Pi(M_k) \leq \Delta\Pi(M_j)$, then $\Delta CS(M_k) < \Delta CS(M_j)$.*

Proof. Suppose instead that $\Delta CS(M_k) \geq \Delta CS(M_j)$. Then there exists a $\overline{c}'_k > \overline{c}_k$ such that $\Delta CS(k, \overline{c}'_k) =$

$\Delta CS(M_j)$. But this implies (using Lemma 2 for the first inequality and Lemma 3 for the second) that $\Delta\Pi(M_k) > \Delta\Pi(k, \overline{c}'_k) > \Delta\Pi(M_j)$, a contradiction. $\square$

3.2 Main Result

We can now turn to the optimal policy of the antitrust authority. Recall that the antitrust authority can without loss restrict itself to approval sets in which the set of acceptable cost levels for a merger between firm 0 and each firm k, $\mathcal{A}_k \subseteq [l, h_k]$, is a union of closed intervals. Throughout we restrict attention to such policies.[5] Let $\overline{a}_k \equiv \max\{\overline{c}_k : \overline{c}_k \in \mathcal{A}_k\}$ denote the largest allowable post-merger cost level for a merger (i.e., the "marginal merger") between firms 0 and k. Also let $\underline{\Delta CS}_k \equiv \Delta CS(k, \overline{a}_k)$ and $\underline{\Delta\Pi}_k \equiv \Delta\Pi(k, \overline{a}_k)$ denote the changes in consumer surplus and bilateral profit, respectively, induced by that marginal merger. These are the lowest levels of consumer surplus and bilateral profit in any allowable merger between firms 0 and k.

At first glance, one may be tempted to conjecture that the antitrust authority can achieve its goal by simply approving any proposed merger that is CS-nondecreasing, i.e., for every $k \geq 1$, setting $\mathcal{A}_k = [l, \overline{a}_k]$, where $\overline{a}_k$ is such that $\Delta CS(k, \overline{a}_k) = 0$. Figure 2(a) illustrates such a policy for a case in which $K = 4$. In the figure, the approval set $\mathcal{A}$ is shown by the heavily-traced sections of the merger curves. In fact, this is not an optimal policy. To see this, suppose the antitrust authority instead adopts an approval policy $\mathcal{A}'$ that imposes a slightly tougher standard on the largest merger: setting $\mathcal{A}'_k = \mathcal{A}_k$ for each merger $k \neq 4$, and setting $\mathcal{A}'_4 = \{M_4 : \Delta CS(M_4) \geq \varepsilon\}$ for $\varepsilon > 0$ sufficiently small. This acceptance set is shown in Figure 2(b). The two policies differ only in the event that the most profitable feasible and acceptable merger under approval policy $\mathcal{A}$, $M^*(\mathfrak{F}, \mathcal{A})$, lies in set $\mathcal{A}\backslash\mathcal{A}'$, i.e., only when $M^*(\mathfrak{F}, \mathcal{A}) = M_4$ and $\Delta CS(M_4) \in [0, \varepsilon)$. Conditional on this event, the expected change in consumer surplus under approval policy $\mathcal{A}$ is bounded from above by ε, which approaches zero as ε becomes small. Under the alternative approval policy $\mathcal{A}'$, and conditioning on the same event, the firms will propose the next-most profitable acceptable merger (which must involve a target $k < 4$). Since the two policies do not differ in their acceptance sets for such smaller mergers, the expected change in consumer surplus under $\mathcal{A}'$ thus converges to $E_{\mathfrak{F}}\left[\Delta CS\left(M^*(\mathfrak{F}\backslash M_4, \mathcal{A})\right) | \Delta\Pi\left(M^*(\mathfrak{F}\backslash M_4, \mathcal{A})\right) \leq \underline{\Delta\Pi}_4\right] > 0$ as ε becomes small. Hence, the expected change in consumer surplus is larger under $\mathcal{A}'$ than under the naive approval policy $\mathcal{A}$.

Since the naive policy of approving any CS-nondecreasing merger is not optimal, how should the antitrust authority construct its approval policy to maximize expected consumer surplus? Our main result is the following:

Proposition 1. *Any optimal approval policy $\mathcal{A}$ approves the smallest merger if and only if it is CS-nondecreasing, approves only mergers $k \in \mathcal{K}^+ \equiv \{1, ..., \widehat{K}\}$ with positive probability ($\widehat{K}$ may equal K), and satisfies $0 = \underline{\Delta CS}_1 < \underline{\Delta CS}_2 < ... < \underline{\Delta CS}_{\widehat{K}}$ for all $k \leq \widehat{K}$. That is, the lowest level of consumer surplus change that is acceptable to the antitrust authority equals zero for the smallest merger M_1, is strictly positive for every other merger M_k with $k > 1$, and is monotonically increasing in the size of the merger, while the largest merger(s) may never be approved.*

According to Lemma 3, there is a systematic misalignment between firms' proposal incentives and the interests of the antitrust authority: firms tend to have an incentive to propose a merger that is

[5] Thus, when we state that any optimal policy must have a particular form, we mean any optimal interval policy of this sort. There are other optimal policies that add or subtract in addition some measure zero sets of mergers, since these have no effect on expected consumer surplus.

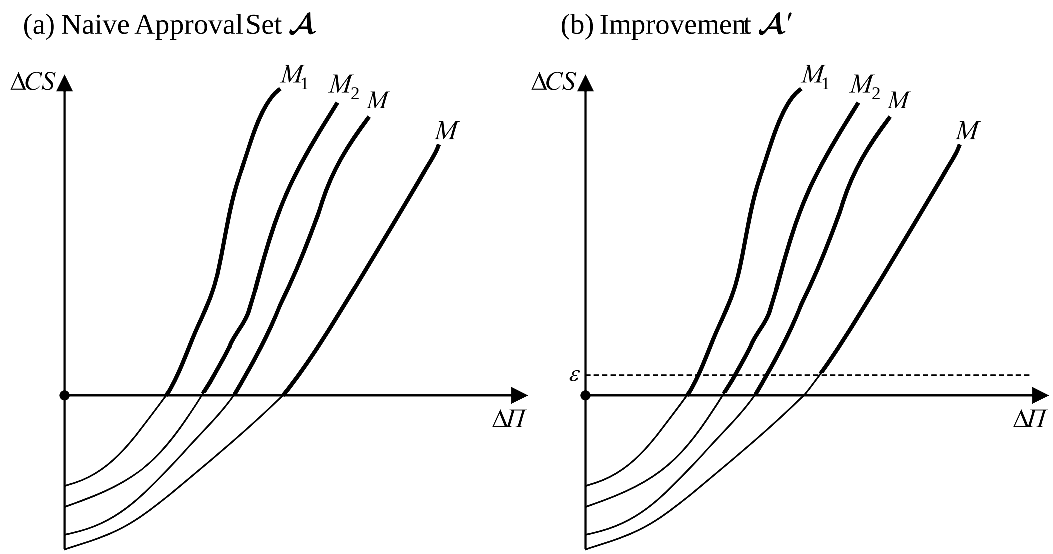

Figure 2: The "naive" policy that accepts all mergers that do not decrease consumer surplus is not optimal. Here, requiring a strictly positive increase in consumer surplus to approve merger M_4 raises expected consumer surplus.

larger (in terms of the target's pre-merger size) than the one that would maximize consumer surplus. Proposition 1 shows that to compensate for this intrinsic bias in firms' proposal incentives, the antitrust authority optimally adopts a higher minimum CS-standard the larger is the proposed merger. Here we give a heuristic derivation of the result; see the formal proof in the Appendix for details. We organize our discussion in "steps" corresponding to those in the formal proof in the Appendix.

Step 1. Observe, first, that the optimal policy $\mathcal{A}$ does not approve CS-decreasing mergers. That is, $\underline{\Delta CS}_k \geq 0$ for all $k \in \mathcal{K}^+$, where $\mathcal{K}^+$ is the set of targets for whom the probability of having a merger $M_k \in \mathcal{A}$ is strictly positive. For any set $\mathcal{A}$ that does approve such mergers, the antitrust authority can increase the expected consumer surplus by instead adopting the alternative policy $\mathcal{A}^+$ that differs from $\mathcal{A}$ only in that it does not contain CS-decreasing mergers. In Figure 3, two such approval sets are depicted in heavy trace. Now, in the event in which, under policy $\mathcal{A}$, the most profitable feasible and allowable merger would have been CS-decreasing, $\Delta CS(M^*(\mathfrak{F},\mathcal{A})) < 0$, the merger that is proposed instead under $\mathcal{A}^+$ is instead CS-nondecreasing. In any other event, the two policies induce the same outcome. Hence, the expected consumer surplus induced by policy $\mathcal{A}$ is lower than that induced by the alternative policy $\mathcal{A}^+$.

Step 2. Next, note that every CS-nondecreasing smallest merger (M_1) must be included in the optimal approval set. If not, as in the set $\mathcal{A}$ shown in Figure 4(a), we could change the approval set $\mathcal{A}$ by adding all CS-nondecreasing mergers M_1, resulting in the alternative approval set $\mathcal{A}'$ shown in Figure 4(b). This change of approval sets matters only in the event in which, under $\mathcal{A}'$, a CS-nondecreasing merger M_1 would be proposed and approved while, under $\mathcal{A}$, this merger would not be approved and firms would therefore propose the next-most profitable allowable merger (which may be the null merger

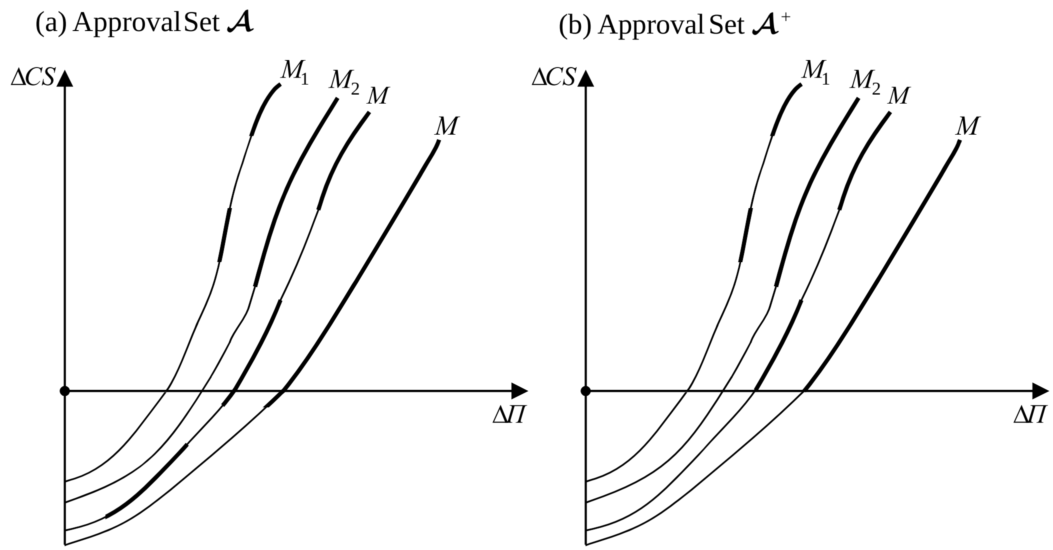

Figure 3: Changing the approval set $\mathcal{A}$ by blocking all mergers that reduce consumer surplus, resulting in approval set $\mathcal{A}^+$, raises expected consumer surplus.

M_0). From Corollary 1, this next-most profitable allowable merger must increase consumer surplus by less than merger M_1. Hence, expected consumer surplus is higher under the alternative approval set $\mathcal{A}'$ than under $\mathcal{A}$.

Step 3. In any optimal approval set $\mathcal{A}$, the consumer surplus level of the marginal merger $M_k = (k, \overline{a}_k)$, $k \in \mathcal{K}^+$, equals the expected CS-level of the next-most profitable acceptable merger, so that $\underline{\Delta CS}_k = E_k^{\mathcal{A}}(\overline{a}_k)$, as illustrated in Figure 5 for $k = 2$, where the expectation $E_2^{\mathcal{A}}(\overline{a}_2)$ is the expected level of ΔCS, conditional on the next-most-profitable merger being in the shaded region. To see this indifference condition, suppose first that the consumer surplus level of the marginal merger M_k is less than the expected CS-level of the next-most profitable acceptable merger, i.e., $\underline{\Delta CS}_k < E_k^{\mathcal{A}}(\overline{a}_k)$. Consider changing the approval set $\mathcal{A}$ by removing all mergers M_k with $\overline{c}_k \in (\overline{a}_k - \varepsilon, \overline{a}_k]$, thereby increasing $\underline{\Delta CS}_k$. For $\varepsilon > 0$ sufficiently small, this change in the approval set increases expected consumer surplus.[6] Similarly, if $\underline{\Delta CS}_k > E_k^{\mathcal{A}}(\overline{a}_k)$, the antitrust authority can increase expected consumer surplus by adding to the approval set all mergers $M_k \in (\overline{a}_k, \overline{a}_k + \varepsilon)$ for $\varepsilon > 0$ sufficiently small.[7]

Step 4. Next, we can see that any optimal approval policy $\mathcal{A}$ has the property that the increase in bilateral profit induced by a marginal merger is greater for larger mergers: $\underline{\Delta\Pi}_j \leq \underline{\Delta\Pi}_k$ for $j < k$, $j, k \in \mathcal{K}^+$. Panel (a) of Figure 6, where $\underline{\Delta\Pi}_2 > \underline{\Delta\Pi}_3$, depicts a situation where this property is not satisfied. Intuitively, the merger $\widehat{M}_2$ directly above the marginal merger $(3, \overline{a}_3)$, has a higher level of ΔCS than does $(3, \overline{a}_3)$, while resulting in the same expected ΔCS if it is rejected. Hence, if $(3, \overline{a}_3)$ is approved, so should be $\widehat{M}_2$, or more precisely, so should those in the set $\overline{\mathcal{A}}_2^{\varepsilon}$ (for small ε) shown in Figure 6(b).

[6] Note that $k \in \mathcal{K}^+$ implies that $\overline{a}_k > l$, so that $\overline{a}_k - \varepsilon > l$ for $\varepsilon > 0$ sufficiently small.
[7] By Step 1 and Assumption 2, we have $\overline{a}_k < h_k$, implying that $\overline{a}_k + \varepsilon < h_k$ for $\varepsilon > 0$ sufficiently small.

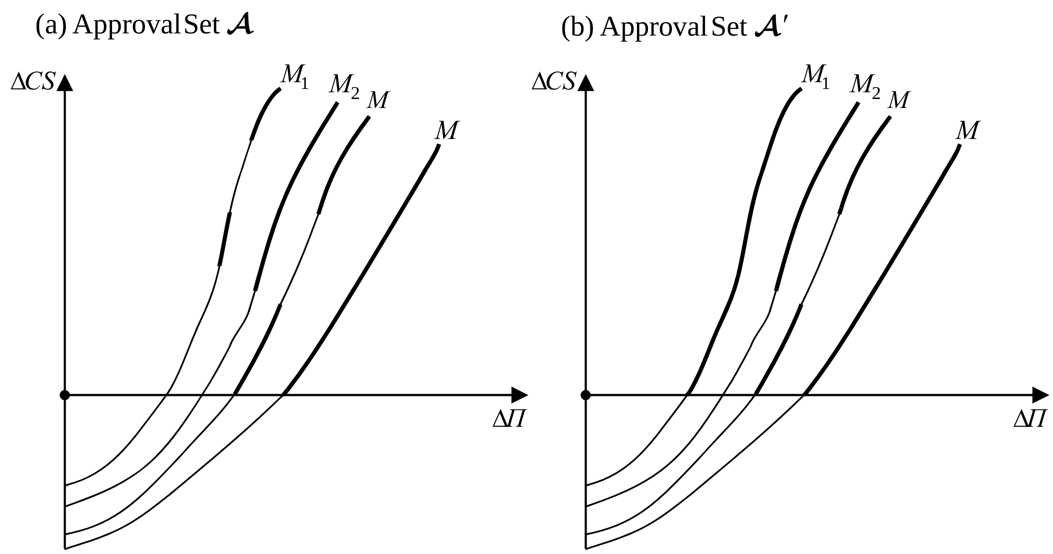

Figure 4: Changing the approval set $\mathcal{A}$ by approving the smallest merger M_1 whenever it does not reduce consumer surplus, resulting in approval set $\mathcal{A}'$, raises expected consumer surplus.

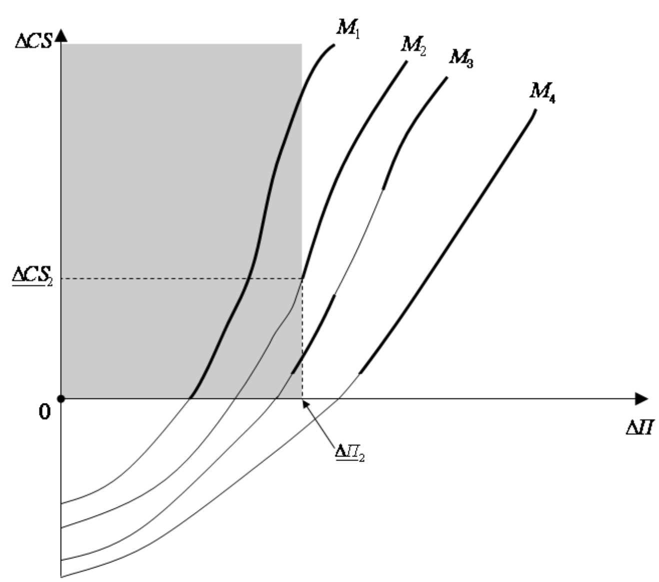

Figure 5: The optimal approval policy is such that the increase in consumer surplus induced by the marginal merger M_k (shown here for k=2) equals the expected consumer surplus change from the next-most-profitable acceptable merger, conditional on the marginal merger being the most profitable merger in the set of feasible and acceptable mergers. The next-most-profitable acceptable merger must therefore lie in the shaded region.

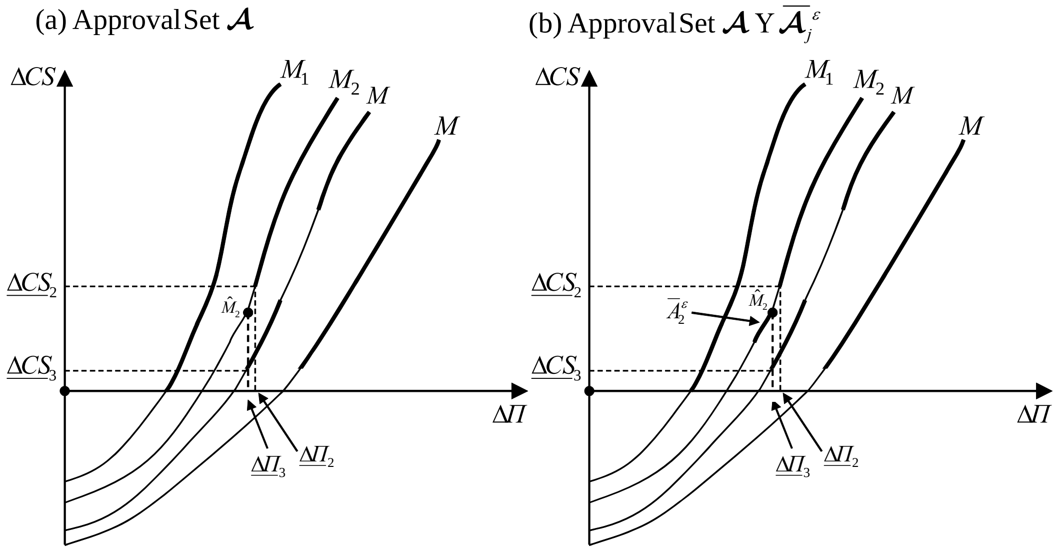

Figure 6: Panel (a) shows a situation where $\underline{\Delta\Pi}_k$ is not increasing in k; panel (b) shows an improvement in the approval set.

Step 5. Next, we can show that in any optimal approval policy $\mathcal{A}$, the consumer surplus increase induced by the marginal merger is strictly greater for larger mergers, i.e., $\underline{\Delta CS}_j < \underline{\Delta CS}_k$ for $j < k$, $j, k \in \mathcal{K}^+$. A situation in which this is not true is illustrated in Figure 7, where $\underline{\Delta CS}_2 \geq \underline{\Delta CS}_3$. By the indifference condition of Step 3, $\underline{\Delta CS}_3$ must equal the expected ΔCS of the next-most profitable allowable merger, i.e., $\underline{\Delta CS}_3 = E_3^{\mathcal{A}}(\overline{a}_3)$. Now, this expectation is the weighted average of the expected ΔCS in two events. First, the next-most profitable allowable merger, say M', may be more profitable than the marginal merger $(2, \overline{a}_2)$, i.e., $\Delta\Pi(M') \in [\underline{\Delta\Pi}_2, \underline{\Delta\Pi}_3)$. In this event, M' must (by Step 4) involve a smaller target (either firm 1 or 2). Hence, the expected ΔCS in this event strictly exceeds $\underline{\Delta CS}_2$. Second, the next-most profitable acceptable merger M' may be less profitable than the marginal merger $(2, \overline{a}_2)$, i.e., $\Delta\Pi(M') < \underline{\Delta\Pi}_2$. By the indifference condition of Step 4, the expected ΔCS in this event is exactly equal to $\underline{\Delta CS}_2$. Taking the weighted average of these two events, we conclude that $\underline{\Delta CS}_3 = E_3^{\mathcal{A}}(\overline{a}_3) > \underline{\Delta CS}_2$, a contradiction.

Step 6. Finally, we argue that if there exists a merger M_j that will never be approved under the optimal policy $\mathcal{A}$, then no larger merger M_k, $k > j$, will ever be approved either: i.e., $k \notin \mathcal{K}^+$ implies $k+1 \notin \mathcal{K}^+$. To see this, observe that $\Delta CS(k, l) > \Delta CS(k+1, l) > \Delta CS(k+1, \overline{a}_{k+1})$ [the first inequality follows because the sum of costs after merger $(k+1, l)$ is lower than that after merger (k, l), whereas the second follows by Lemma 2], so as in Step 5, there is a nonmonotonicity in the ΔCS-levels of the marginal mergers with firms k and $k+1$: i.e., $\Delta CS(k, l) > \Delta CS(k+1, \overline{a}_{k+1})$. The result then follows using an argument like that in Step 5; see the Appendix for details.

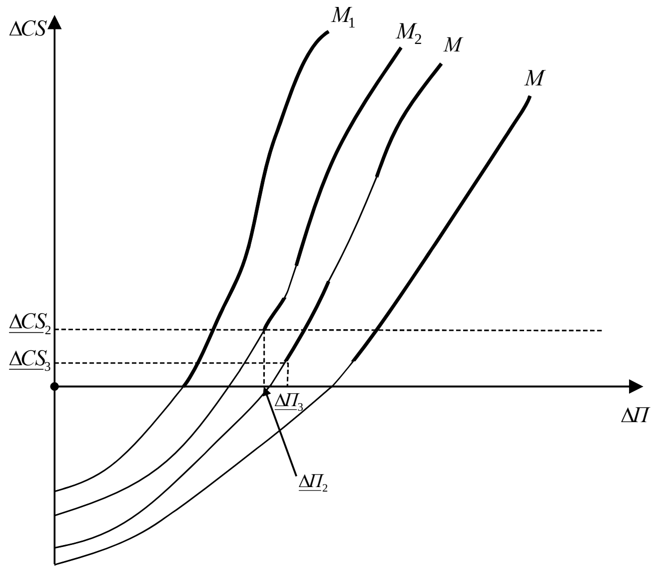

Figure 7: The optimal approval set is such that the consumer surplus increase induced by the marginal merger M_j, is less than that by the marginal larger merger M_k, $k > j$, i.e., $\underline{\Delta CS}_j < \underline{\Delta CS}_k$. In the figure, $\underline{\Delta CS}_2 > \underline{\Delta CS}_3$, which is a violation of that property.

4 Cutoff Policies and Comparative Statics

Proposition 1 shows that in any optimal policy the least efficient acceptable merger involving a target k [the marginal merger $M_k = (k, \overline{a}_k)$] involves a larger increase in consumer surplus (and larger increase in bilateral profit) the larger is the target. Moreover, the result holds for any distributions of post-merger marginal costs. However, it does not fully characterize those marginal mergers. Indeed, while we know that the marginal merger $M_k = (k, \overline{a}_k)$ satisfies the indifference condition $\Delta CS(M_k) = E_k^{\mathcal{A}}(\overline{a}_k)$, the expectation $E_k^{\mathcal{A}}(\overline{a}_k)$ depends on the acceptance sets for mergers other than k (i.e, on $\mathcal{A}_j$, $j \neq k$), whose optimal forms depend in turn on merger k's acceptance set $\mathcal{A}_k$.

Identifying the marginal merger for each target would be much simpler if we knew that the optimal policy had a "cutoff" structure, in which, for each target k, any mergers with greater efficiencies than the marginal merger are accepted. Specifically, a cutoff policy $\mathcal{A}^C$ is defined by a set of marginal cost cutoffs, $(\overline{a}_1^C, ..., \overline{a}_K^C)$, such that $M_k = (k, \overline{c}_k) \in \mathcal{A}^C$ if and only if $\overline{c}_k \leq \overline{a}_k^C$. In that case, Proposition 1 would imply that the marginal mergers could be found by a simple recursive procedure: accept all CS-nondecreasing mergers M_1 [i.e., set $\overline{a}_1^C = \widehat{c}_1(Q^0)$], then for $k = 2, ..., K$ recursively identify the largest post-merger cost level $\overline{a}_k^C$ for which $\Delta CS_k(k, \overline{a}_k^C) = E_k^{\mathcal{A}^C}(\overline{a}_k^C)$, where now the expectation $E_k^{\mathcal{A}^C}(\overline{a}_k^C)$ depends only on the already-determined cutoffs for mergers $1, ..., k-1$. If $\Delta CS(k, \overline{c}_k) < E_k^{\mathcal{A}^C}(\overline{c}_k)$ for all $\overline{c}_k \in [l, h_k]$, then no such cutoff exists for merger M_k, so that $\mathcal{A}_k^C = \varnothing$. Moreover, this will also imply that $\mathcal{A}_{k'}^C = \varnothing$ for all $k' > k$.

Unfortunately, however, as the following example illustrates, the optimal policy need not have a cutoff structure. (For simplicity, the example considers a case where, contrary to the assumption of

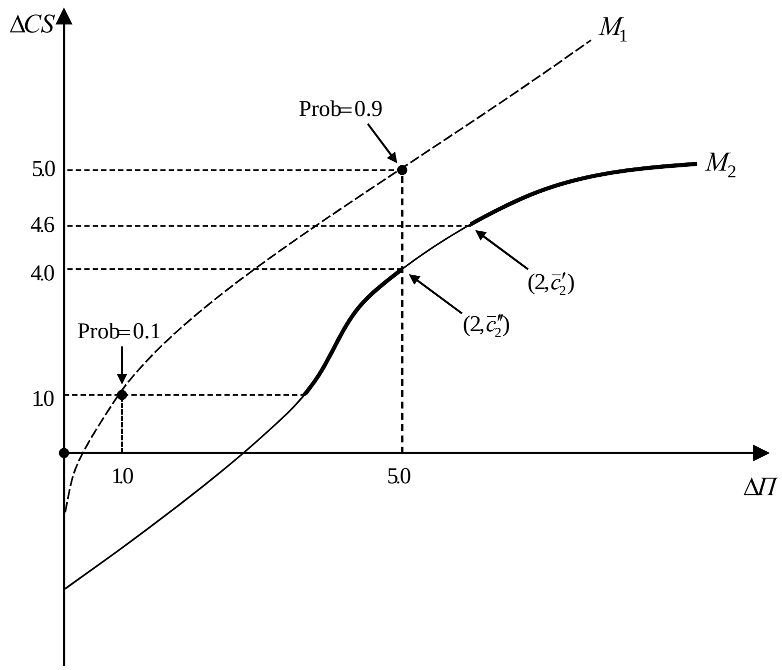

Figure 8: The figure depicts an example where the optimal approval set does not have a cutoff structure.

the model, one of the mergers has a finite support of post-merger marginal costs. But the same insight would obtain if we perturbed the example and assumed that the support is continuous with no atoms.)

Example 1. Suppose that there are two possible mergers, M_1 and M_2 . The smaller merger, M_1, is always feasible. Its post-merger marginal cost is either $\overline{c}_1 = l$ or $\overline{c}_1 = h_1$, where the probability on the latter is 0.9. The corresponding changes in consumer surplus and bilateral profit are given by $(\Delta\Pi(1,l), \Delta CS(1,l)) = (5,5)$ and $(\Delta\Pi(1,h_1), \Delta CS(1,h_1)) = (1,1)$. The unconditional expected increase in consumer surplus from approving M_1 is thus equal to 4.6. The post-merger marginal cost of the larger merger, M_2, has a continuous support $[l, h_2]$ with no atoms, satisfying $\Delta CS(2,h_2) < 1$ and $\Delta CS(2,l) > 5$. Let $\overline{c}_2'$ be such that $\Delta CS(2,\overline{c}_2') = 4.6$ and $\overline{c}_2''$ be such that $\Delta\Pi(2,\overline{c}_2'') = 5$, and assume that $\overline{c}_2' < \overline{c}_2''$. It is straightforward to verify that, in this case, the optimal approval policy $\mathcal{A}^*$ is such that $\mathcal{A}_1^* = \{l, h_1\}$ and $\mathcal{A}_2^* = [l, \overline{c}_2'] \cup [\overline{c}_2'', \overline{a}_2]$. This situation is illustrated in Figure 8. To see why the optimal approval policy for M_2 does not have a cut-off structure, note that for any post-merger marginal cost $\overline{c}_2 \in (\overline{c}_2', \overline{c}_2'')$, M_2 would always be the proposed merger if it were approved when proposed. But the induced change in consumer surplus from M_2 would be less than 4.6, which is the expected change in consumer surplus from M_1. The optimal policy corrects for this bias in firms' proposal policies by rejecting merger M_2 whenever $\overline{c}_2 \in (\overline{c}_2', \overline{c}_2'')$.

Nonetheless, our next result provides a sufficient condition that ensures that the recursively-defined cutoff policy is in fact optimal. To proceed, let $\mathcal{A}^C(J) \subseteq \Pi_{k\in J}[l, h_k]$ denote the recursively-defined cutoff policy when only mergers with targets in set J are possible; that is, when we suppose that there is no possibility for a merger with any target $k \notin J$. [The policy $\mathcal{A}^C(J)$ specifies $\#J$ cutoffs.] For convenience, when $J = \mathcal{K}$ we write $\mathcal{A}^C \equiv \mathcal{A}^C(\mathcal{K})$. We also let $\overline{a}_k^C(J)$ denote the cutoff level of marginal cost for a merger with target k in cutoff policy $\mathcal{A}^C(J)$.

In addition, for a set of targets $J \subseteq \mathcal{K}$, define the realized set of feasible mergers to be $\mathfrak{F}_J$, and the function

$$ECS(\overline{\Delta\Pi}; \mathcal{A}, J) \equiv E_{\mathfrak{F}_J}[\Delta CS(M^*(\mathfrak{F}_J, \mathcal{A}))|\Delta\Pi(M^*(\mathfrak{F}_J, \mathcal{A})) \leq \overline{\Delta\Pi}]$$

as the expected value of ΔCS under policy $\mathcal{A} \subseteq \Pi_{k\in J}[l, h_k]$ from the most profitable acceptable merger involving targets in set J, conditional on that merger's increase in bilateral profit being no greater than $\overline{\Delta\Pi}$.[8] Note that the structure of $\mathcal{A}$ at profit levels above $\overline{\Delta\Pi}$ affects the value of this conditional expectation by changing the conditional distributions of post-merger marginal costs. Specifically, the probability of a merger in set $\mathcal{M}_j \subseteq \{M_j : \Delta\Pi(M_j) \leq \overline{\Delta\Pi}\}$ being feasible conditional on the most profitable acceptable merger having a profit level below $\overline{\Delta\Pi}$ is $\Pr(M_j \in \mathcal{M}_j) \times [1- \Pr(\Delta\Pi(M_j) > \overline{\Delta\Pi}$ and $M_j \in \mathcal{A}_j)]^{-1}$. Note that an optimal policy $\mathcal{A}^* \subseteq \Pi_{k\in\mathcal{K}}[l, h_k]$ is an element of $\arg\max_{\mathcal{A}} ECS(\infty; \mathcal{A}, \mathcal{K})$.

We then have:

Proposition 2. *Suppose that for every $J \subseteq \mathcal{K}$ with $1 \in J$ the following property holds:*[9]

Every merger $M_k = (k, \overline{c}_k) \in \mathcal{A}^C(J)$ with $\overline{c}_k < \overline{a}_k^C(J)$ has $\Delta CS(M_k) > ECS(\Delta\Pi(M_k); \mathcal{A}^C(J\backslash k), J\backslash k)$. (9)

Then, the cutoff policy $\mathcal{A}^C$ is an optimal policy.

Proof. In the Appendix. □

While Proposition 2 does not offer a condition on primitives, it allows us to verify that the recursively-derived cutoff policy is optimal. The following example provides an illustration of its use.

Example 2. Consider a four-firm industry (so $N = 4$) in which firm 0 can merge with each of the other firms (so $K = 3$). Industry inverse demand is $P(Q) = 1 - Q$. Pre-merger marginal costs are $c_0 = c_2 = 0.5$, $c_1 = 0.55$, and $c_3 = 0.45$, so the pre-merger market shares are $s_0 = s_2 = 1/4$, $s_1 = 1/8$, and $s_4 = 3/8$. The naive policy marginal cost cutoffs (where any CS-nondecreasing merger is accepted) are $\overline{a}_1 = 0.45$, $\overline{a}_2 = 0.40$, $\overline{a}_3 = 0.35$. Now suppose that each merger has a 3/4 probability of being feasible (so $\theta_k = 0.75$ for $k = 1, 2, 3$) and that, conditional on being feasible, the post-merger marginal cost is distributed with a beta distribution between the merger's naive cutoff and 0.2.[10] When $\beta = 1$ and $\alpha = 5$, expected consumer surplus could be increased by 6.44% if there were there no informational asymmetry between the firms and the antitrust authority (so the merger that increased

[8] Thus, $E_k^{\mathcal{A}}(\overline{c}_k) = ECS(\Delta\Pi(k, \overline{c}_k); \mathcal{A}_{\mathcal{K}\backslash k}, \mathcal{K}\backslash k)$ where $\mathcal{A}_{\mathcal{K}\backslash k} \equiv \Pi_{j\in\mathcal{K}\backslash k}\mathcal{A}_j$.

[9] Note that property (9) necessarily holds for $j = 1$; the assumption made here is that it holds for all $j > 1$.

[10] One can think of this situation as having a 1/4 probability of there being no CS-increasing merger, and a 3/4 probability of a CS-increasing merger. The beta distribution has a pdf $f(x|\alpha, \beta)$ that is proportional to $x^{\alpha-1}(1-x)^{\beta-1}$. Its mean is the lower bound of its support plus a fraction $\alpha/(\alpha + \beta)$ of the difference between its support's upper and lower bounds. When $\beta = 1$ and $\alpha > 1$, as in the cases we study here, the pdf is an increasing function, so that small efficiency gains are more likely than large ones. The lower bound of 0.2 is chosen to ensure that all firms remain active after any merger.

consumer surplus most would always be implemented). In this setting, one can verify that the sufficient condition of Proposition 2 is satisfied, so the recursively-defined cutoff policy is optimal. The cutoffs in this optimal policy are $\overline{a}_1 = 0.45$, $\overline{a}_2 = 0.383$, and $\overline{a}_3 = 0.316$, with associated changes in consumer surplus of $\underline{\Delta CS}_1 = 0$, $\underline{\Delta CS}_2 = .00170$, and $\underline{\Delta CS}_3 = .00346$. The optimal policy achieves 90.30% of the first-best increase in expected consumer surplus, while the naive policy achieves 79.83% of this amount. This outcome is shown in Table 1. The table also shows the results when $\alpha = 3$ and $\alpha = 1$ ($\alpha = 1$ is a uniform distribution). Both of these cases also satisfy the sufficient condition in Proposition 2. As α decreases, the distributions of post-merger marginal costs (conditional on the merger being feasible) shift towards lower costs and the expected consumer surplus gain increases. However, the gain from the fully optimal policy relative to the naive policy falls.

Table 1:

α	$\underline{CS}_1$	$\underline{CS}_2$	$\underline{CS}_3$	first-best: % gain in $E[CS]$	naive policy: % of first-best gain	optimal policy: % of first-best gain
5	0	0.00170	0.00346	6.44	79.83	90.30
3	0	0.00170	0.00457	9.55	87.34	92.13
1	0	0.00099	0.00571	18.15	93.67	94.10

4.1 Comparative Statics

When cutoff rules are optimal we can explore how changes in underlying parameters alter the nature of the optimal policy. We provide two such results here, assuming that the optimal policy has a cutoff structure. Consider, first, changes in the feasibility probabilities. Intuitively, lower feasibility probabilities should move the optimal policy toward the naive one. For example, as all θ's approach zero, the optimal policy approaches the naive policy, since there is almost no chance that two mergers are feasible. Our first result, which builds on this intuition, examines the effect of a decrease in the likelihood that a merger with a given target k is feasible.

Proposition 3. *Consider an decrease in the probability of merger M_k's feasibility from θ_k to $\theta'_k < \theta_k$, assuming that M_k is initially approved with positive probability (i.e., $k \leq \widehat{K}$). Then, under the optimal merger approval policy, $\underline{\Delta CS}_j{}' = \underline{\Delta CS}_j$ for any weakly smaller merger M_j, $j \leq k$, and $\underline{\Delta CS}'_j < \underline{\Delta CS}_j$ for any larger merger M_j, $j > k$, that is approved with positive probability.*

Proof. In the Appendix. □

Our second result concerns a change in pre-merger costs:

Proposition 4. *Consider a reduction in firm 0's marginal cost from c_0 to $c'_0 < c_0$. Under the optimal merger approval policy, this induces a decrease in all post-merger marginal cost cutoffs: $\overline{a}'_k < \overline{a}_k$ for every $1 \leq k \leq \widehat{K}$.*

Proof. In the Appendix. □

5 Extensions

In this section, we consider five extensions of our baseline model. First, we consider alternative bargaining processes among firms. Second, we analyze the optimal merger approval policy in a more general setting by relaxing two assumptions: (i) every merger involves two firms, and (ii) firm 0 is party to any merger. Third, adopting an aggregative game approach, we consider the case of price competition with differentiated products (CES and multinomial logit demand structures). Fourth, we study the optimal merger approval policy when the antitrust authority cares not only about consumer surplus but also about producer surplus. Finally, we extend the model by allowing for synergies in fixed costs.

5.1 Other Bargaining Processes

In our analysis so far, we have focused on the case where the bargaining process between firms is given by the offer game [Segal (1999)]. In the offer game, firm 0 makes a take-it-or-leave-it offer to a target of its choosing and is therefore able to extract all of the gain in bilateral profit. The equilibrium of the offer game therefore results in the proposal of the merger that maximizes the change in the bilateral profit of the merger partners in the realized set of feasible and acceptable mergers.

It is straightforward to see that the same outcome would obtain if the bargaining power were more evenly distributed between firms, provided firm 0 can extract the same fixed fraction of the gain in bilateral profit with each target. This would hold, for example, if firm 0 first selects a potential target, say firm k, and then the bargaining process between firm 0 and firm k is given by the alternating offer bargaining game of Rubinstein (1982), assuming that all potential targets have the same discount factor.

In the following, we explore two alternative bargaining processes. First, we consider the benchmark case of efficient bargaining. Second, we consider the case where there is efficient bargaining only among a subset of firms (including all of the firms that are involved in potential mergers). We show that, in both cases, our main result continues to hold: the optimal approval policy has the property that the minimum CS-standard is increasing in the size of the proposed merger.

5.1.1 Efficient Bargaining

Suppose the outcome of the bargaining process is efficient for the firms in the industry in the sense that it maximizes aggregate profit. That is, we assume that, from the realized set of feasible and acceptable mergers, $\mathfrak{F} \cap \mathcal{A}$, firms choose to propose merger

$$M^* (\mathfrak{F}, \mathcal{A}) \equiv \arg \max_{M_k \in (\mathfrak{F} \cap \mathcal{A})} \Delta\Pi(M_k),$$

where $\Delta\Pi(M_k)$ now denotes the change in *aggregate profit* induced by merger M_k,

$$\Delta\Pi(M_k) \equiv \sum_{i \in \mathcal{N} \setminus \{0\}} \pi_i(M_k) - \sum_{i \in \mathcal{N}} \pi_i^0.$$

There are several bargaining processes that would lead to aggregate profit maximization:

1. "Coasian bargaining" among all firms under complete information.

2. A "menu auction" in which each firm $i \neq 0$ submits a nonnegative bid $b_i(M_k) \geq 0$ to firm 0 for each merger $M_k \in (\mathfrak{F} \cap \mathcal{A})$, $k \geq 1$, and firm 0 then selects the merger that maximizes its profit, inclusive of these bids. [Firm 0's profit from selecting the null merger M_0 is $\pi_0(M_0)$.] Bernheim and Whinston (1996) show that there is an efficient equilibrium which, in this setting, implements the merger that maximizes aggregate profit.

3. The target (firm 0) committing to a sales mechanism. Jehiel, Moldovanu, and Stacchetti (1996) show that an optimal mechanism has the following structure in our setting: Firm 0 proposes to implement the aggregate profit-maximizing merger $M^*(\mathfrak{F},\mathcal{A})$ and requires the payment $\pi_i(M^*(\mathfrak{F},\mathcal{A}))-\pi_i(\underline{M}_i)$ from each firm $i \neq 0$, where $\underline{M}_i$ is the merger in set $(\mathfrak{F}\cap\mathcal{A})\backslash M_i$ that minimizes firm i's profit. If a firm i does not accept participation in the mechanism when all other firms do, then the principal commits to proposing merger $\underline{M}_i$ to the antitrust authority.[11] Given this mechanism, there is an equilibrium in which all firms participate in the mechanism, and merger $M^*(\mathfrak{F},\mathcal{A})$ is proposed.[12]

We claim that Proposition 1 continues to hold when bargaining is efficient. The key steps in the argument are the following: First, note that Lemma 1 states that a CS-neutral merger M_k, $k \geq 1$, raises not only the bilateral profit of the merger partners but also aggregate profit, i.e., that $\Delta\Pi(M_k) > 0$.

Lemma 2, however, does not extend to the case of aggregate profit without imposing an additional condition. We therefore assume that a reduction in post-merger marginal cost increases aggregate profit if the merger is CS-nondecreasing:

Assumption 3. *If merger M_k, $k \geq 2$, is CS-nondecreasing [i.e., if $\overline{c}_k \leq \widehat{c}(Q^0)$], then reducing its post-merger marginal cost $\overline{c}_k$ increases the aggregate profit $\sum_{i\in\mathcal{N}\backslash\{0\}}\pi_i(M_k)$.*[13]

In fact, this assumption must hold for merger M_k if pre-merger cost differences are small enough so that the sum of the pre-merger market share of firms 0 and k weakly exceeds the pre-merger share of any other firm, i.e., $s_0^0 + s_k^0 \geq \max_{j\neq 0,k} s_j^0$.[14] To see why Assumption 3 holds in this case, note that summing up the first-order conditions for profit maximization following merger M_k [conditions (2) and (3)] yields

$$\begin{aligned}\sum_{i\in\mathcal{N}\backslash\{0\}}\pi_i(M_k) &= \sum_{i\in\mathcal{N}\backslash\{0,k\}}\left[P(Q(M_k))-c_i\right]q_i(M_k)+\left[P(Q(M_k))-\overline{c}_k\right]q_k(M_k) \\ &= \left|[Q(M_k)]^2\,P'(Q(M_k))\right|H(M_k), \qquad (10)\end{aligned}$$

where $H(M_k) \equiv \sum_{i\in\mathcal{N}\backslash\{0\}}(s_i(M_k))^2$ is the post-merger industry Herfindahl index. Assumption 1 ensures that the first term, $\left|Q^2P'(Q)\right|$, is increasing in Q. By Lemma 2, a reduction in post-merger marginal cost $\overline{c}_k$ leads to a larger $Q(M_k)$, so a sufficient condition for the claim to hold is that reducing the merged firm's marginal cost $\overline{c}_k$ induces an increase in $H(M_k)$. Under Assumption 1, a decrease in the merged firm's marginal cost $\overline{c}_k$ increases the share of the merged firm and decreases the share of every other firm. Since $s_0^0 + s_k^0 \geq \max_{j\neq 0,k} s_j^0$ implies $s_k(M_k) \geq \max_{j\neq 0,k} s_j(M_k)$ for any CS-nondecreasing merger M_k with $k \in \{1,...,K\}$, this induced change in market shares increases the post-merger Herfindahl index $H(M_k)$ (see Lemma 7 in the Appendix).

[11] Similar to Bernheim and Whinston's (1996) menu auction, firms $i \neq 0$ make payments even when they are not party to a merger.

[12] To see that firm 0 wants to propose merger $M^*(\mathfrak{F},\mathcal{A})$, note that using this type of mechanism its optimal merger proposal solves

$$\max_{M_k\in(\mathfrak{F}\cap\mathcal{A})}\Pi(M_k)-\sum_{i\neq 0}\pi_i(\underline{M}_i),$$

which is equivalent to $\max_{M_k\in(\mathfrak{F}\cap\mathcal{A})}\Pi(M_k)$.

[13] Note that we make this assumption only for mergers with targets $k \geq 2$ because the arguments in Proposition 1 rely on monotonicity of the merger curves only for mergers other than the smallest merger.

[14] In Section 5.2, where we consider more general sets of mergers, we provide a weaker sufficient condition for Assumption 3 to hold.

Next, the systematic misalignment of interests between firms and the antitrust authority, as stated in Lemma 3, is also present when bargaining is efficient:

Lemma 4. *Suppose two mergers, M_j and M_k, with $j < k$, induce the same non-negative change in consumer surplus, $\Delta CS(M_j) = \Delta CS(M_k) \geq 0$. Then, the larger merger M_k induces a greater increase in aggregate profit: i.e., $\Delta\Pi(M_k) > \Delta\Pi(M_j) > 0$.*

Proof. In the Appendix. □

Finally, given Assumption 3 and Lemma 4, we can draw a figure just like Figure 1, but with $\Delta\Pi$ representing the aggregate profit arising from a merger. As a result, all of the steps in the proof of Proposition 1 continue to hold with efficient bargaining.

5.1.2 Efficient Bargaining Between a Subset of Firms

Suppose instead that the outcome of the bargaining process maximizes the joint profit of only a subset of firms, $\mathcal{L}$, that includes firm 0 and all of the targets, i.e., $(\{0\} \cup \mathcal{K}) \subseteq \mathcal{L} \subset \mathcal{N}$. That is, the proposal rule is

$$M^*(\mathfrak{F}, \mathcal{A}) \equiv \arg \max_{M_k \in (\mathfrak{F} \cap \mathcal{A})} \Delta\Pi(M_k),$$

where $\Delta\Pi(M_k)$ now denotes the induced change in the joint profit of the firms in set $\mathcal{L}$, $\Delta\Pi(M_k) \equiv \sum_{i \in \mathcal{L} \setminus \{0\}} \pi_i(M_k) - \sum_{i \in \mathcal{L}} \pi_i^0$.

Under the same conditions as in the case of efficient bargaining, Proposition 1 carries over to this bargaining process. The key point is the following: If any CS-nondecreasing merger or any reduction in a merged firm's marginal cost induces an increase in aggregate profit, then it also induces an increase in the joint profit of the firms in set $\mathcal{L}$. This follows because both a CS-nondecreasing merger and a reduction in a firm's post-merger marginal cost weakly reduce the profit of any nonmerging firm, including the firm(s) not in set $\mathcal{L}$. This observation has several implications. First, it means that part (iv) of Lemma 1 continues to hold if we replace aggregate profit by the joint profit of the firms in set $\mathcal{L}$. Second, it also means that Assumption 3 implies that a reduction in the post-merger marginal cost $\overline{c}_k$ raises the joint profit of the firms in set $\mathcal{L}$ for any CS-nondecreasing merger. Third, Lemma 4 continues to hold if we replace the induced change in aggregate profit by the induced change in the joint profit of the firms in $\mathcal{L}$. This follows because the two mergers in the statement of the lemma, M_j and M_k, induce (by assumption) the same change in consumer surplus, so the profit of any firm $i \neq j, k$ is the same under both mergers. As a result, we can again draw a figure like Figure 1, and all of the steps in the proof of Proposition 1 carry over to this case.

5.2 General Sets of Mergers

So far, we have assumed that there is a single firm, firm 0, that is part of every potential merger. Moreover, we have assumed that every merger involves only two firms, firm 0 and one target. In this section, we relax both of these assumptions by allowing for general sets of mergers. As the offer game no longer seems an appropriate bargaining process once there is no single firm that is party to every potential merger, we focus on efficient bargaining. We continue to assume that at most one merger can be proposed to the antitrust authority. We provide sufficient conditions under which the main result of the paper carries over to this more general setting. In particular, we show that the key criterion according to which the antitrust authority should optimally discriminate between alternative mergers is the naively-computed post-merger Herfindahl index. This naively-computed post-merger index is frequently used by antitrust authorities in merger analysis as it is entirely based on readily available information on pre-merger market structure.

To proceed, let $m_k \geq 2$ denote the number of merger partners in merger M_k and let $\overline{c}_{M_k}$ denote the realized post-merger marginal cost of merger M_k. It is straightforward to see that the characterization of CS-neutral mergers in Lemma 1 extends to any $m_k \geq 2$. In particular, any CS-neutral merger raises aggregate profit. In Section 5.1, we have shown that aggregate profit following merger M_k is proportional to the post-merger Herfindahl index $H(M_k)$, where the proportionality factor depends only on the post-merger aggregate output $Q(M_k)$ [see (10)]. Observe that for a CS-*neutral* merger M_k, Lemma 1 implies that the actual post-merger Herfindahl index equals the naively-computed index:

$$\begin{aligned} H(M_k) &\equiv [s_k(M_k)]^2 + \sum_{i \notin M_k} [s_i(M_k)]^2 \\ &= \left[\sum_{i \in M_k} s_i^0 \right]^2 + \sum_{i \notin M_k} \left[s_i^0 \right]^2 \equiv H^{naive}(M_k). \end{aligned}$$

Thus, for any two CS-neutral mergers M_j and M_k, regardless of the number of merger partners, the merger that induces a greater naively-computed post-merger Herfindahl index also induces a greater increase in aggregate profit:

$$H^{naive}(M_k) > H^{naive}(M_j) \Leftrightarrow \Delta\Pi(M_k) > \Delta\Pi(M_j).$$

Hence, provided that merger curves slope upward in the positive orthant of $(\Delta\Pi, \Delta CS)$-space and do not intersect, Proposition 1 carries over to this more general setting, where a "larger" merger now refers to a merger that induces a greater increase in the naively-computed post-merger Herfindahl index.

Under what conditions do the curves for CS-nondecreasing mergers slope upward and not intersect? To identify such conditions, we first observe that the inverse of the slope of the curve for merger M_k in $(\Delta\Pi, \Delta CS)$-space is given by (see Lemma 8 in the Appendix)

$$\frac{d\Delta\Pi(M_k)}{d\Delta CS(M_k)} = -2 - \left[\frac{P''(Q(M_k))Q(M_k)}{P'(Q(M_k))} \right] H(M_k) + \left[\frac{2}{P'(Q(M_k))Q(M_k)} \right] \left[\frac{r(Q(M_k); \overline{c}_{M_k})}{dQ(M_k)/d\overline{c}_{M_k}} \right], \quad (11)$$

where $r(Q;c)$ is again the cumulative best reply of a firm with marginal cost c to aggregate output Q.

We now use expression (11) to identify condtions under which the merger curves are upward-sloping and non-intersecting.[15] As earlier, merger curves are upward-sloping in the positive orthant whenever the pre-merger joint market share of any merging firms exceed the pre-merger share of the largest nonmerging firm. However, expression (11) allows us to derive a weaker condition than this:[16]

[15] Condition (11) also offers an alternative method to establish Lemma 4. To see this, observe that, in our baseline model, if two mergers induce the same change in consumer surplus, ΔCS, and the same change in aggregate profit, $\Delta\Pi$, then the two mergers also induce the same aggregate output Q and the same post-merger Herfindahl index H. Moreover, in our baseline model, the firm resulting from a larger merger has a larger output $r(Q; \overline{c}_M)$ (as it faces a larger $\sum_{i \neq M} c_i$, and so must have a lower $\overline{c}_M$ if it induces an equal CS-level). Hence, (11) implies in that model that if there were a point of intersection, the curve of the larger merger would have a larger value of $d\Delta\Pi/d\Delta CS$, hence a flatter curve, which yields a contradiction since the larger merger's curve must cross from below at the first crossing since the larger merger's curve lies further to the right where $\Delta CS = 0$.

[16] That condition (12) below holds when $s_{M_k}^{naive} \geq \max_{i \notin M_k} s_i^0$ follows from the facts that in this case $H^{naive}(M_k) \leq s_{M_k}^{naive}$ and $(N - m_k + 2)s_{M_k}^{naive} \geq 1$. (Note that, in general, the Herfindahl index is bounded above by the share of the largest firm.)

Lemma 5. *The merger curve of merger M_k slopes upward in the positive orthant of* $(\Delta\Pi, \Delta CS)$*-space if the merged firm's naively-computed post-merger market share* $s_{M_k}^{naive} \equiv \sum_{i \in M_k} s_i^0$ *and the naively-computed post-merger Herfindahl index* $H^{naive}(M_k)$ *satisfy*

$$s_{M_k}^{naive} \geq \frac{H^{naive}(M_k)}{2} \geq 1 - (N - m_k + 2)s_{M_k}^{naive}, \tag{12}$$

where $N+1$ *is the pre-merger number of firms (and thus* $N - m_k + 2$ *is the number of firms following merger* M_k*).*

Proof. In the Appendix. □

Now consider when the merger curves are non-intersecting. We will use expression (11) to provide conditions under which two merger curves cannot cross; that is, their ranking must be the same as their ranking where $\Delta CS = 0$. We will show this by contradiction, showing that the curve further to the right at $\Delta CS = 0$ must have a smaller slope wherever the two curves cross. Since aggregate profit and consumer surplus are the same wherever the curves cross, so must be the industry Herfindahl index and aggregate quantity. By (11), this means the slopes at that point are ordered by the values of $r(Q(M_k); \overline{c}_{M_k})/(dQ(M_k)/d\overline{c}_{M_k})$. The following lemma provides a condition under which those quantities are ordered in the correct way to give us an analog of Lemma 3:

Lemma 6. *Consider two mergers* M_j *and* M_k*, with* $m_j \geq m_k$*. If the firms in* M_k *jointly produce more pre-merger than the firms in* M_j *(i.e.,* $\sum_{i \in M_k} s_i^0 > \sum_{i \in M_j} s_i^0$*) and if the naively-computed post-merger Herfindahl index is larger when* M_k *is implemented than when* M_j *is implemented [i.e.,* $H^{naive}(M_k) > H^{naive}(M_j)$*], then the curve relating to merger* M_k *lies to the right of that relating to merger* M_j *in the positive orthant of* $(\Delta\Pi, \Delta CS)$*-space.*

Proof. In the Appendix. □

Finally, if all mergers have the same minimum of the support of post-merger marginal costs, denoted l, the maximum CS-increase that the smaller merger can achieve is larger than that of the larger merger.[17] Hence, under the assumptions of Lemmas 5 and 6, the merger curves have all of the properties required to obtain our main result, the analog of Proposition 1.

For example, one special case in which this result can be applied arises where there are three potential mergers, one involving firms 1 and 2, a second involving firms 1 and 3, and a third involving firms 2 and 3. As before, the three mergers are mutually exclusive but, in contrast to the baseline model, there is no longer a single firm that is party to every potential merger. In this case, the two conditions of Lemma 6 are satisfied if the mergers have the same ranking by both the product and the sum of the merging firms' pre-merger market shares.

[17] To see this, consider a larger merger M_k and a smaller merger M_j, $j < k$. The maximum ΔCS induced by the larger merger M_k is $\Delta CS(k, l)$. The assumption in Lemma 6 that the firms in M_k produce more pre-merger implies that $\Sigma_{i \in M_k} c_i < \Sigma_{i \in M_j} c_i$. Since agrgegate quantity depends only on the sum of firms' costs, this in turn implies that if the two mergers induce the same change in consumer surplus, then $\overline{c}_{M_k} < \overline{c}_{M_j}$. Thus, $\Delta CS(k, l) = \Delta CS(j, \overline{c}_{M_j})$ implies that $\overline{c}_{M_j} > l$. Since $\Delta CS(j, \overline{c}_{M_j})$ is decreasing in $\overline{c}_{M_j}$, the maximum CS-increase for the smaller merger, $\Delta CS(j, l)$, must be larger than that of the larger merger: i.e., $\Delta CS(j, l) > \Delta CS(k, l)$.

5.3 Differentiated Products

In our analysis we have assumed that firms produce a homogeneous good and compete in a Cournot fashion. Restricting attention to the case of efficient bargaining between firms and mergers between firm 0 and a single target firm $k \in \mathcal{K}$, we now show that our main insights carry over to the case where firms compete in prices and produce symmetrically differentiated goods with consumers having CES or multinomial logit demand. Specifically, we assume that the initial market structure is such that every firm produces one differentiated good. If a merger M_k is proposed and approved, then a merged firm produces the two products of its merger partners at an identical post-merger marginal costs, $\overline{c}_k$.

CES Demand. In the CES model, the utility function of the representative consumer is given by

$$U = \left(\sum_{i=0}^{N} X_i^{\rho} \right)^{1/\rho} Z^{\alpha},$$

where $\rho \in (0,1)$ and $\alpha > 0$ are parameters, X_i is consumption of differentiated good i, and Z is consumption of the numeraire. Utility maximization implies that the representative consumer spends a constant fraction $1/(1+\alpha)$ of his income Y on the $N+1$ differentiated goods (and the remainder on the numeraire). Using the normalization $Y/(1+\alpha) \equiv 1$, the resulting demand for differentiated good i is

$$X_i = \frac{p_i^{-\lambda-1}}{\sum_{j=0}^{N} p_j^{-\lambda}},$$

where p_i is the price of good i, and $\lambda \equiv \rho/(1-\rho)$. The consumer's indirect utility can be written as

$$V = (1+\alpha) \ln Y + \frac{1}{\lambda} \ln \left(\sum_{j=0}^{N} p_j^{-\lambda} \right). \tag{13}$$

Multinomial Logit Demand. In the multinomial logit model, expected demand for product i is given by

$$X_i = \frac{\exp\left(\frac{a-p_i}{\mu}\right)}{\sum_{j=0}^{N} \exp\left(\frac{a-p_j}{\mu}\right)},$$

where $a > 0$ and $\mu > 0$ are parameters, and p_j the price of product j. Letting Y denote income, the indirect utility of the representative consumer can be written as

$$V = Y + \mu \ln \left[\sum_{j=0}^{N} \exp\left(\frac{a-p_j}{\mu} \right) \right]. \tag{14}$$

The CES and multinomial logit models share important features with the Cournot model. In particular, all of these models can be written as "aggregative games." That is, the profit a firm obtains from its plant or product i can be written as

$$\pi(\psi_i, c_i; \Psi),$$

where $\psi_i \geq 0$ is the firm's strategic variable, c_i is its (constant) marginal cost, and $\Psi \equiv \sum_j \psi_j$ is an aggregator summarizing the "aggregate outcome." [If a merged firm runs two plants or produces two products at the same marginal cost $\overline{c}_k$ and chooses the same value ψ_k of its strategic variable for both of its plants or products, then its total profit is $2\pi(\psi_k, \overline{c}_k; \Psi)$.] Further, consumer surplus is an increasing function of the aggregator, and does not depend on its composition, so that it can be written as $V(\Psi)$.

In the Cournot model, ψ_i is output q_i and Ψ is aggregate output Q, so that profit can be written as $\pi(\psi_i, c_i; \Psi) = (P(\Psi) - c_i)\psi_i$ and consumer surplus as $V(\Psi) = \int_0^{\Psi}(P(x) - P(\Psi))dx$. In the CES model, we have $\psi_i = p_i^{-\lambda}$ and $\Psi = \sum_j p_j^{-\lambda}$, so that profit from product i can be written as

$$\pi(\psi_i, c_i; \Psi) = [\psi_i^{-1/\lambda} - c_i]\frac{\psi_i^{(\lambda+1)/\lambda}}{\Psi}.$$

From the indirect utility (13), it follows that consumer surplus is an increasing function of Ψ. Finally, in the multinomial logit model, we have $\psi_i = \exp((a - p_i)/\mu)$ and $\Psi = \sum_j \exp((a - p_j)/\mu)$, so that profit from product i can be written as

$$\pi(\psi_i, c_i; \Psi) = [a - \mu \ln \psi_i - c_i]\,\frac{\psi_i}{\Psi}.$$

From the indirect utility (14), it follows that consumer surplus is an increasing function of Ψ.

In the Appendix, we show that the equilibrium profit functions of these three models share some important properties. Using this common structure, we show in the Appendix that if merger M_k is CS-neutral, then it raises the joint profit of the merging firms as well as aggregate profit. Moreover, a reduction in post-merger marginal cost increases the merged firm's profit and, provided pre-merger differences between firms are not too large, aggregate profit. Moreover, if any two mergers M_j and M_k, $k > j$, induce the same nonnegative change in consumer surplus, then the larger merger M_k induces a greater increase in aggregate profit than the smaller merger M_j. In sum, in the two differentiated goods models, the merger curves have the same features in $(\Delta CS, \Delta\Pi)$-space as in the Cournot model. Our main result, Proposition 1, therefore carries over as well.

5.4 Alternative Welfare Standard

In our baseline model, we have assumed that the antitrust authority seeks to maximize consumer surplus. While this is in line with the legal standard in the U.S., the EU, and many other countries, it might seem unsatisfactory that the antitrust authority completely ignores any effect of its policy on producer surplus. We now show that our main result extends to the case where the antitrust authority seeks to maximize any convex combination of consumer surplus and aggregate surplus. We focus on the case of efficient bargaining between firms, but discuss the offer game at the end of the section.

Specifically, suppose the antitrust authority's welfare criterion is $W \equiv CS + \lambda\Pi$, where $\lambda \in [0, 1]$. When $\lambda = 1$, welfare W thus amounts to aggregate surplus. Let

$$\Delta W(M_k) \equiv \Delta CS(M_k) + \lambda\Delta\Pi(M_k)$$

denote the change in welfare induced by approving merger M_k. We will say that merger M_k is W-increasing [resp. W-decreasing] if $\Delta W(M_k) > 0$ [resp. $\Delta W(M_k) < 0$]. It is W-nondecreasing [resp. W-nonincreasing] if $\Delta W(M_k) \geq 0$ [resp. $\Delta W(M_k) \leq 0$].

Since a W-increasing merger may be CS-decreasing, we require a slightly stronger version of Assumption 3:

Assumption 3' *If merger M_k for $k \geq 2$ is W-nondecreasing, then (i) reducing its post-merger marginal cost $\overline{c}_k$ increases the aggregate profit, and (ii) $\overline{c}_k < \min\{c_0, c_k\}$; i.e., the merger involves synergies.*

Assumption 3' must hold if pre-merger marginal cost differences are sufficiently small. To see this, consider the extreme case where all firms have the same pre-merger marginal cost c. Then, for merger M_k to be W-nondecreasing, it must involve synergies in that $\overline{c}_k < c$.[18] Hence, if M_k is W-nondecreasing, the merged firm is the firm with the lowest marginal cost post merger. Reducing the merged firm's marginal cost $\overline{c}_k$ therefore increases aggregate output Q (thereby raising $|Q^2 P'(Q)|$) and the Herfindahl index H. From equation (10), a lower level of post-merger marginal cost $\overline{c}_k$ thus results in a greater aggregate profit. By continuity of consumer and producer surplus in marginal costs, it follows that if pre-merger marginal cost differences are sufficiently small, then $\Delta W(M_k) \geq 0$ implies that $\overline{c}_k < \min\{c_0, c_k\}$ and aggregate profit is decreasing in $\overline{c}_k$.

We also impose the following analog of Assumption 2:

Assumption 2' *For all $k \in \mathcal{K}$, the support of the post-merger cost distribution includes both W-increasing and W-nonincreasing mergers: i.e., $\Delta W(k, h_k) \leq 0 < \Delta W(k, l)$.*

Assumption 3' allows us to obtain a slightly stronger version of Lemma 4:

Lemma 4' *Suppose two W-nondecreasing mergers, M_j and M_k, with $k > j \geq 1$, induce the same change in consumer surplus, $\Delta CS(M_j) = \Delta CS(M_k)$. Then the larger merger M_k induces a greater increase in aggregate profit: i.e., $\Delta\Pi(M_k) > \Delta\Pi(M_j) > 0$.*

Proof. In the Appendix. □

Figure 9(a) depicts the merger curves in $(\Delta\Pi, \Delta CS)$-space. The downward-sloping lines are isowelfare curves, each with slope $-\lambda$; the dashed line is the isowelfare curve corresponding to no welfare change, $\Delta W = 0$. Lemma 4' states that, above the dashed line $\Delta W = 0$, the curve corresponding to a larger merger lies everywhere to the right of that corresponding to a smaller merger.

Figure 9(b) depicts the merger curves in $(\Delta\Pi, \Delta W)$-space. Note that each merger curve has a positive horizontal intercept: since a CS-nondecreasing merger increases aggregate profit, a W-neutral merger must be CS-decreasing and therefore increase aggregate profit. Moreover, each curve is upward-sloping in the positive orthant. Finally, in the positive orthant, the curve of a larger merger lies everywhere to the right of that of a smaller merger.

Letting $\underline{\Delta W}_k \equiv \Delta W(k, \overline{a}_k)$ denote the welfare level of the "marginal merger," i.e., the lowest welfare level in any allowable merger between firms 0 and k, we can then establish the following analog of our main result (Proposition 1) for the case in which the antitrust authority maximizes an arbitrary convex combination of consumer surplus and aggregate surplus:

Proposition 1' *Any optimal approval policy $\mathcal{A}$ approves the smallest merger if and only if it is W-nondecreasing, and satisfies $0 = \underline{\Delta W}_1 < \underline{\Delta W}_j < \underline{\Delta W}_k$ for all $j, k \in \mathcal{K}^+$, $1 < j < k$, where $\mathcal{K}^+ \subseteq \mathcal{K}$ is the set of mergers that is approved with positive probability.*

[18] To see this, suppose otherwise that $\overline{c}_k \geq c$. We can decompose the induced change in market structure into two steps: (i) a move from $N+1$ to N firms, each with marginal cost c, and (ii) an increase in the marginal cost of one firm from c to $\overline{c}_k \geq c$. Step (i) induces a reduction in aggregate output but does not affect average production costs, and so reduces W. Step (ii) weakly reduces aggregate output and weakly increases average costs in the industry, and so weakly reduces W.

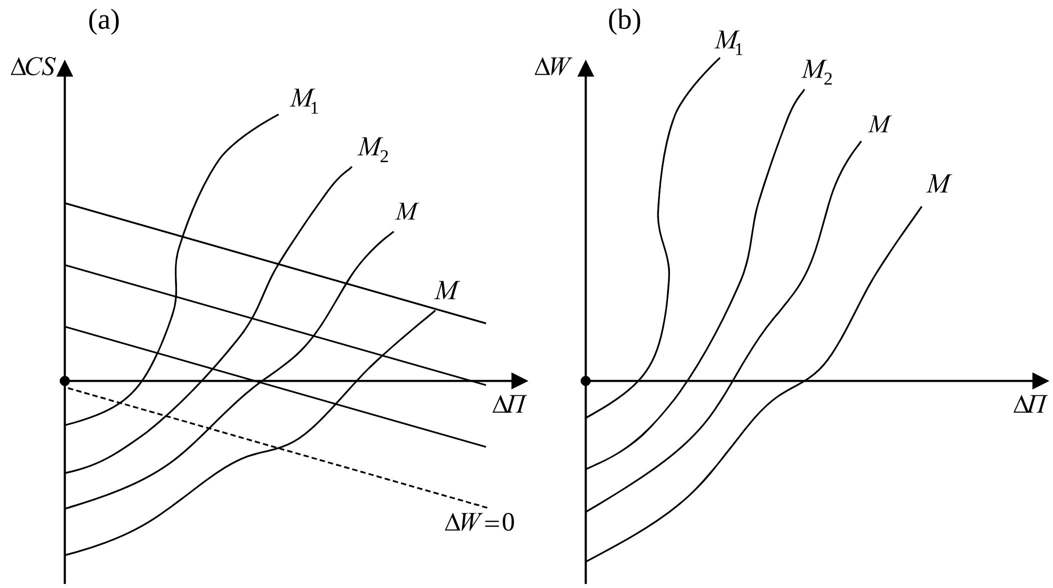

Figure 9: Panel (a) shows the merger curves in $(\Delta\Pi, \Delta CS)$-space. The downward-sloping lines are the iso-welfare curves. Panel (b) shows the merger curves in $(\Delta W, \Delta CS)$-space.

Proposition 1' differs from Proposition 1 only in that we can no longer show that if a merger with target k is never approved, then neither is a merger with any larger target.[19]

Establishing a parallel result for the case of the offer game is more difficult. However, the result extends provided the antitrust authority does not put too much weight on aggregate profit (i.e., provided $\lambda > 0$ is sufficiently small). Intuitively, this follows because our merger-curve graph is affected continuously by changes in λ, so for values near 0 we are very close to the situation in Figure 1.

5.5 Synergies in Fixed Costs

So far, we have assumed that firms have constant returns, implying that all merger-specific efficiencies involve marginal cost savings. We now consider the case where firms have to incur fixed costs, part of which may be saved by merging, and identify conditions under which our main result carries over to this setting. The discussion that follows applies to our baseline (offer-game) model as well to the other bargaining models discussed in Section 5.1, with $\Delta\Pi$ appropriately reinterpreted.

Let f_i denote the fixed cost of firm i and asume that it is small enough that firm i remains active following any merger by other firms. A feasible merger M_k is now described by $M_k = (k, \overline{c}_k, \overline{f}_k)$, where $\overline{f}_k \in [\overline{f}^l_k, \overline{f}^h_k] \subset \mathbb{R}_+$ is the realization of its post-merger fixed cost. The merger induces a fixed cost saving if $f_0 + f_k - \overline{f}_k \equiv \alpha_k > 0$. Graphically, a fixed cost saving shifts the merger curve in a parallel fashion (by the amount of the saving) to the right in $(\Delta\Pi, \Delta CS)$-space. Thus, the possibility

[19] The reason is that Step 6 in the proof of Proposition 1 does not carry over, as we cannot guarantee that $\Delta W(k, l) > \Delta W(k+1, l)$. However, the same type of argument as in Step 6 can be used to show that if $j \notin \mathcal{K}^+$, $k \in \mathcal{K}^+$, and $j < k$, then $\Delta W(j, l) < \underline{\Delta W}_k$. That is, if a merger with target k is never approved, then any larger merger that is approved must have a greater increase in consumer surplus than the most efficient possible merger M_k.

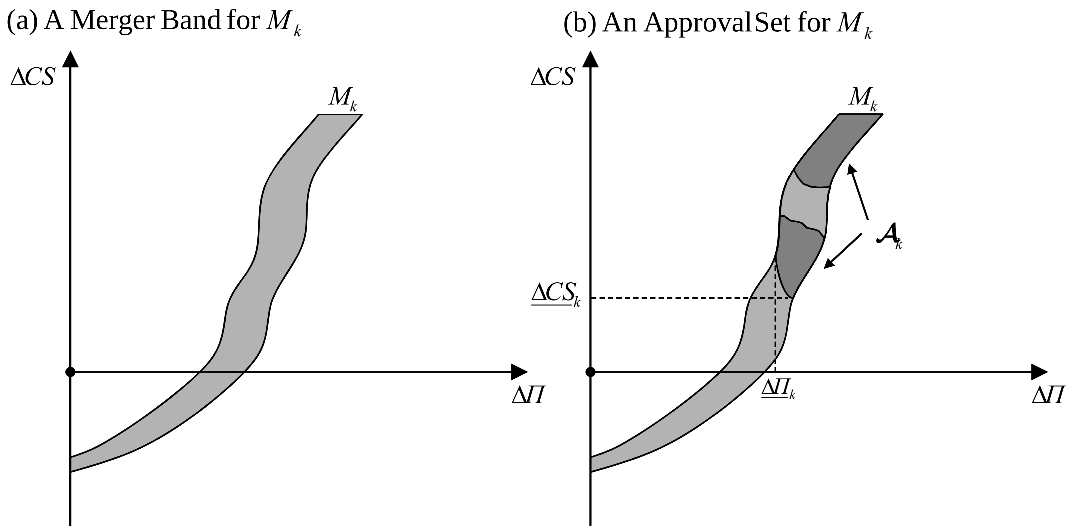

Figure 10: The figure depicts merger bands when mergers create both marginal and fixed cost savings in panel (a) and a possible approval set in panel (b).

of fixed cost savings implies that the merger curves in $(\Delta\Pi, \Delta CS)$-space are now "broad bands," with each point in the band of merger M_k corresponding to a different realization of $(\overline{c}_k, \overline{f}_k)$, and with the horizontal width of the band given by $\left|\overline{f}_k^h - \overline{f}_k^l\right|$ at any $\Delta CS(M_k)$. Figure 10 depicts the merger band for merger M_k.

When a feasible merger is proposed, the antitrust authority can observe all aspects of that merger, including the induced fixed cost saving. The antitrust authority's approval set is now described by $\mathcal{A} \equiv \left\{M_k : \left(\overline{c}_k, \overline{f}_k\right) \in \mathcal{A}_k\right\} \cup M_0$, where $\mathcal{A}_k \subseteq [l, h_k] \times [\overline{f}_k^l, \overline{f}_k^h]$. Without loss of generality, we restrict attention to approval sets that are *regular* in the sense that every $\mathcal{A}_k$ is the closure of its interior, i.e., $\mathcal{A}_k = \operatorname{cl}(\operatorname{int}(\mathcal{A}_k))$. Let $\overline{a}_k(\overline{f}_k) \equiv \max\{\overline{c}_k : (\overline{c}_k, \overline{f}_k) \in \mathcal{A}_k\}$ denote the largest allowable post-merger marginal cost level for a merger between firms 0 and k, conditional on the realized post-merger fixed cost $\overline{f}_k$. Let $\underline{\Delta CS}_k(\overline{f}_k) \equiv \Delta CS(k, \overline{a}_k(\overline{f}_k), \overline{f}_k)$ and $\underline{\Delta\Pi}_k(\overline{f}_k) \equiv \Delta\Pi(k, \overline{a}_k(\overline{f}_k), \overline{f}_k)$ denote the changes in consumer surplus and bilateral profits, respectively, induced by the "marginal merger" between firms 0 and k given $\overline{f}_k$, and let $\underline{\Delta CS}_k \equiv \min_{\overline{f}_k \in [\overline{f}_k^l, \overline{f}_k^h]} \underline{\Delta CS}_k(\overline{f}_k)$ and $\underline{\Delta\Pi}_k \equiv \min_{\overline{f}_k \in [\overline{f}_k^l, \overline{f}_k^h]} \underline{\Delta\Pi}_k(\overline{f}_k)$ denote the lowest levels of ΔCS and $\Delta\Pi$, respectively, in any acceptable merger M_k. Figure 10(b) depicts an approval set for merger M_k and shows $\underline{\Delta CS}_k$ and $\underline{\Delta\Pi}_k$.

An immediate observation is the following. Suppose that fixed cost savings are nonnegative and perfectly correlated across mergers, so that $\alpha_k = \alpha \geq 0$ for every feasible merger $M_k \in \mathfrak{F}$. Then the optimal approval set is constant in α in the sense that $(\overline{c}_k, f_0 + f_k - \alpha) \in \mathcal{A}_k$ if and only if $(\overline{c}_k, f_0 + f_k - \alpha') \in \mathcal{A}_k$, from which it follows that $\underline{\Delta CS}_k(\overline{f}_k) = \underline{\Delta CS}_k$ for all $\overline{f}_k$ and k. Moreover, as before, the optimal policy for any α is characterized by Proposition 1. To see this, note that the expected CS-maximizing antitrust authority cares about fixed cost savings only insofar as they affect firms' merger proposals. But if fixed cost savings are perfectly correlated and nonnegative, the profit ranking of mergers (and teh profitability of CS-nondecreasing mergers) is unaffected by the fixed cost realization and all CS-nondecreasing mergers remain profitable.

Suppose now that the realized fixed cost saving of merger M_k can be decomposed as follows:

$$\alpha_k = \alpha + \eta_k,$$

where $\alpha \in [\alpha^l, \alpha^h]$ is the (random or deterministic) component that is common across all feasible mergers (as above) and $\eta_k \in [\eta_k^l, \eta_k^h]$ is the component idiosyncratic to merger M_k. We assume that both the idiosyncratic shocks and post-merger marginal cost realizations are independent across mergers conditional on α, have full support, and no mass points. We assume as well that when merger M_k is proposed, the antitrust authority can observe α and η_k separately (and condition the approval set on both components separately).[20] Using the same arguments as above, it is straightforward to show that the optimal approval set is constant in α. Therefore, for notational simplicity, we will from now on assume that there is no common component (i.e., $\alpha \equiv 0$), so that $\overline{f}_k = f_0 + f_k - \eta_k$.

In the remainder of this section, we assume that $\left|\overline{f}_k^h - \overline{f}_k^l\right|$ is sufficiently small so that the bands of the different mergers are non-overlapping in the positive orthant, as depicted in Figure 11. Thus, if any two mergers M_j and M_k, $j < k$, induce the same nonnegative change in consumer surplus, then the larger merger is more profitable, regardless of the realized fixed cost savings. As fixed cost savings are nonnegative by assumption, the conclusion of Lemma 1 – that a CS-neutral merger is profitable – continues to hold.

Our main result, Proposition 1, carries over to this setting:

Proposition 5. *In the model with fixed cost savings, any optimal approval policy $\mathcal{A}$ approves the smallest merger if and only if it is CS-nondecreasing, approves only mergers $k \in \mathcal{K}^+ \equiv \{1, ..., \widehat{K}\}$ with positive probability ($\widehat{K}$ may equal K), and satisfies $0 = \underline{\Delta CS}_1 < \underline{\Delta CS}_2 < ... < \underline{\Delta CS}_{\widehat{K}}$ for all $k \leq \widehat{K}$.*

Proof. In the Appendix. □

Thus, provided that idiosyncratic fixed cost synergies are small enough that merger bands do not overlap, it remains optimal to adopt a more stringent consumer surplus test for larger mergers. The restriction on the size of fixed cost synergies contrasts with the model of Armstrong and Vickers (2010). Their model, applied to the merger problem, assumes that the distribution of possible $(\Delta\Pi, \Delta CS)$ pairs are the same for each merger and has a rectangular support. An interesting open question is how projects that are ex ante asymmetric in terms of their distribution of $(\Delta\Pi, \Delta CS)$ pairs should be differentially treated when their supports overlap or even coincide.

6 Conclusion

In this paper, we have analyzed the optimal policy of an antitrust authority towards horizontal mergers when there are several mutually exclusive merger possibilities and firms can choose *which* merger to propose to the antitrust authority. In our baseline model, there is a single pivotal firm, firm 0, that can merge with one of several, ex ante heterogeneous merger partners. The merger that is proposed is the result of a simple bargaining process, the "offer game." While the feasibility and post-merger marginal costs of the various potential mergers is stochastic and not known to the antitrust authority,

[20] That is, a feasible merger M_k is described by $M_k = (k, \overline{c}_k, \alpha, \overline{f}_k)$, and the approval set by $\mathcal{A} \equiv \{M_k : (\overline{c}_k, \alpha, \overline{f}_k) \in \mathcal{A}_k\} \cup M_0$, where $\mathcal{A}_k \subseteq [l, h_k] \times [\alpha^l, \alpha^h] \times [\overline{f}_k^l, \overline{f}_k^h]$.

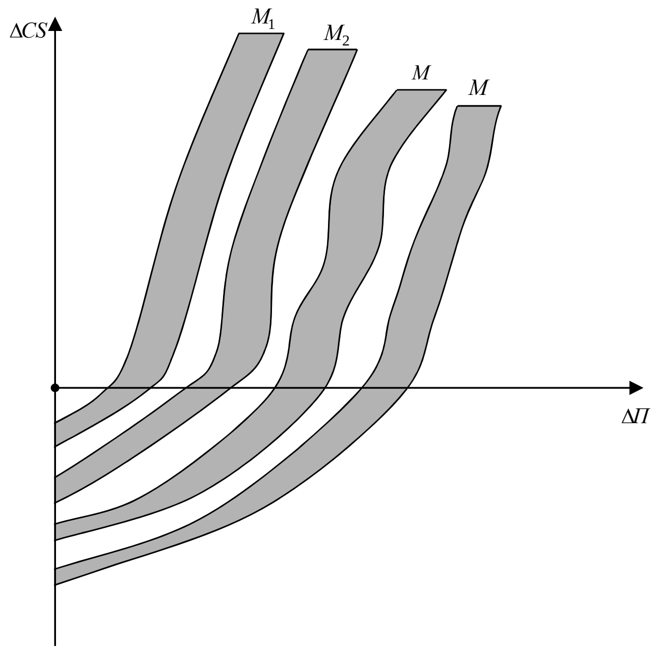

Figure 11: The figure shows merger bands for various possible mergers when mergers can create both fixed and marginal cost synergies.

the antitrust authority can observe the characteristics of the proposed merger. We have shown that the antitrust authority optimally commits to a policy that imposes a tougher standard on mergers that cause a larger increase in the naively-computed Herfindahl index (which is equivalent to a larger pre-merger market share in our baseline model): the required minimum increase in consumer surplus is greater for mergers that are larger in this sense.

We have also seen that our result extends to some other bargaining models, some cases in which the set of possible mergers is richer than the simple one acquierer/multiple target structure of our baseline model, models of differentiated price competition, welfare measures that count aggregate profit, and situations in which merger synergies may arise in fixed costs. Even with these extensions, perhaps the greatest limitation of our model lies in the limited set of possible mergers and bargaining processes that we could handle. The challenge here lies in the small number of tractable models of bargaining with externalities that currently exist in the literature for cases with a rich structure of possible agreements. Further progress on such bargaining models seems critical to gain additional insight into the problem of optimal merger policy.

7 Appendix

7.1 Proofs

Proof of Proposition 1. The proof proceeds in a number of steps.

Step 1. We observe first that an optimal policy does not approve CS-decreasing mergers. That is, $\underline{\Delta CS}_k \geq 0$ for all $k \in \mathcal{K}^+$, where $\mathcal{K}^+$ denotes those targets for whom the probability of having a merger $M_k \in \mathcal{A}$ is strictly positive. To see this, suppose the approval set $\mathcal{A}$ includes CS-decreasing mergers, and consider the set $\mathcal{A}^+ \subseteq \mathcal{A}$ that removes any mergers in $\mathcal{A}$ that reduce consumer surplus. Figure 3 depicts such a pair of approval sets, each containing the points shown with heavy trace. Since this change only matters when the bilateral profit-maximizing merger $M^*(\mathfrak{F},\mathcal{A})$ under set $\mathcal{A}$ is no longer approved under $\mathcal{A}^+$, the change in expected consumer surplus from this change in the approval policy equals $\Pr(M^*(\mathfrak{F},\mathcal{A}) \in \mathcal{A}\backslash\mathcal{A}^+)$, the probability of this event happening, times the conditional expectation

$$E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F},\mathcal{A}^+)) - \Delta CS(M^*(\mathfrak{F},\mathcal{A}))|M^*(\mathfrak{F},\mathcal{A}) \in \mathcal{A}\backslash\mathcal{A}^+].$$

Since $\Delta CS(M^*(\mathfrak{F},\mathcal{A}^+))$ is necessarily nonnegative by construction of $\mathcal{A}^+$, and $\Delta CS(M^*(\mathfrak{F},\mathcal{A}))$ is strictly negative whenever $M^*(\mathfrak{F},\mathcal{A}) \in \mathcal{A}\backslash\mathcal{A}^+$, this change is strictly positive.

Step 2. Next, any smallest merger M_1 that is CS-nondecreasing must be approved. To see this, suppose that the approval set is $\mathcal{A}$ but that $\mathcal{A} \subset \mathcal{A}' \equiv (\mathcal{A} \cup \{(1, \overline{c}_1) : \Delta CS(1, \overline{c}_1) \geq 0\})$. Figure 4 depicts two such sets, $\mathcal{A}$ and $\mathcal{A}'$. Because a change from $\mathcal{A}'$ to $\mathcal{A}$ matters only when the bilateral profit-maximizing merger $M^*(\mathfrak{F},\mathcal{A}')$ under $\mathcal{A}'$ is no longer approved under $\mathcal{A}$, the change in expected consumer surplus by using $\mathcal{A}'$ rather than $\mathcal{A}$ equals $\Pr(M^*(\mathfrak{F},\mathcal{A}') \in \mathcal{A}'\backslash\mathcal{A})$ times

$$E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F},\mathcal{A}')) - \Delta CS(M^*(\mathfrak{F},\mathcal{A}))|M^*(\mathfrak{F},\mathcal{A}') \in \mathcal{A}'\backslash\mathcal{A}]. \tag{15}$$

By Corollary 1 and the fact that $\mathcal{A}'\backslash\mathcal{A}$ contains only smallest mergers (between firms 0 and 1), whenever $M^*(\mathfrak{F},\mathcal{A}') \in \mathcal{A}'\backslash\mathcal{A}$ [which implies $\Delta\Pi(M^*(\mathfrak{F},\mathcal{A}')) > \Delta\Pi(M^*(\mathfrak{F},\mathcal{A}))$] we have $\Delta CS(M^*(\mathfrak{F},\mathcal{A}')) > \Delta CS(M^*(\mathfrak{F},\mathcal{A}))$, so (15) is strictly positive. This can be seen in Figure 4. This implies in particular that $\underline{\Delta CS}_1 = 0$.

Step 3. Next, we claim that in any optimal policy, for all $k \in \mathcal{K}^+$, $\underline{\Delta CS}_k$ must equal the expected change in consumer surplus from the next-most-profitable merger (i.e., from the merger with the second-highest bilateral profit change) $M^*(\mathfrak{F}\backslash(k,\overline{a}_k),\mathcal{A})$, conditional on merger $M_k = (k, \overline{a}_k)$ being the most profitable merger in $\mathfrak{F}\cap\mathcal{A}$. Defining the expected change in consumer surplus from the next-most-profitable merger $M^*(\mathfrak{F}\backslash M_k,\mathcal{A})$, conditional on merger $M_k = (k, \overline{c}_k)$ being the most profitable merger in $\mathfrak{F}\cap\mathcal{A}$, to be

$$\begin{aligned} E_k^{\mathcal{A}}(\overline{c}_k) \quad &\equiv \quad E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F}\backslash M_k,\mathcal{A}))|M_k = (k, \overline{c}_k) \text{ and } M_k = M^*(\mathfrak{F},\mathcal{A})] && (16) \\ &= \quad E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F}\backslash M_k,\mathcal{A}))|M_k = (k, \overline{c}_k) \text{ and } \Delta\Pi(M^*(\mathfrak{F}\backslash M_k,\mathcal{A})) \leq \Delta\Pi(M_k)], && (17) \end{aligned}$$

this means that

$$\underline{\Delta CS}_k = E_k^{\mathcal{A}}(\overline{a}_k). \tag{18}$$

In Figure 5 the possible locations of the next-most-profitable merger when the most profitable merger is $M_2 = (2, \overline{a}_2)$ are shown as a shaded set. The quantity $E_2^{\mathcal{A}}(\overline{a}_2)$ is the expectation of the change in consumer surplus for the merger that has the largest change in bilateral profit among mergers other than M_2, conditional on all of these other mergers lying in the shaded region of the figure.

To see that (18) must hold for all $k \in \mathcal{K}^+$, suppose first that $\underline{\Delta CS}_{k'} > E_{k'}^{\mathcal{A}}(\overline{a}_{k'})$ for some $k' \in \mathcal{K}^+$ and consider the alternative approval set $\mathcal{A} \cup \mathcal{A}_{k'}^{\varepsilon}$ where

$$\mathcal{A}_{k'}^{\varepsilon} \equiv \{M_k : M_k = (k', \overline{c}_{k'}) \text{ with } \overline{c}_{k'} \in (\overline{a}_{k'}, \overline{a}_{k'} + \epsilon)\}.$$

(By Step 1 and Assumption 2, we have $\overline{a}_{k'} < h_{k'}$, implying that $\overline{a}_{k'} + \varepsilon < h_{k'}$ for $\varepsilon > 0$ sufficiently small.) For any $\varepsilon > 0$, the change in expected consumer surplus from changing the approval set from

$\mathcal{A}$ to $\mathcal{A} \cup \mathcal{A}^{\varepsilon}_{k'}$ equals $\Pr(M^*(\mathfrak{F},\mathcal{A} \cup \mathcal{A}^{\varepsilon}_{k'}) \in \mathcal{A}^{\varepsilon}_{k'})$ times

$$E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F},\mathcal{A} \cup \mathcal{A}^{\varepsilon}_{k'})) - \Delta CS(M^*(\mathfrak{F},\mathcal{A}))|M^*(\mathfrak{F},\mathcal{A} \cup \mathcal{A}^{\varepsilon}_{k'}) \in \mathcal{A}^{\varepsilon}_{k'}]. \tag{19}$$

This conditional expectation can be rewritten as

$$E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F},\mathcal{A} \cup \mathcal{A}^{\varepsilon}_{k'})) - E^{\mathcal{A}}_{k'}(\overline{c}_{k'})|M^*(\mathfrak{F},\mathcal{A} \cup \mathcal{A}^{\varepsilon}_{k'}) \in \mathcal{A}^{\varepsilon}_{k'}], \tag{20}$$

where $\overline{c}_{k'}$ is the realized cost level in the bilateral profit-maximizing merger $M^*(\mathfrak{F},\mathcal{A} \cup \mathcal{A}^{\varepsilon}_{k'})$, which is a merger of firms 0 and k' when the conditioning statement is satisfied. By continuity of $\Delta CS(k', \overline{c}_{k'})$ and $E^{\mathcal{A}}_{k}(\overline{c}_{k'})$ in $\overline{c}_{k'}$, there exists an $\overline{\varepsilon} > 0$ such that $\Delta CS(M_{k'}) > E^{\mathcal{A}}_{k}(\overline{c}_{k'})$ for all $M_{k'} \in \mathcal{A}^{\varepsilon}_{k'}$ provided $\varepsilon \in (0, \overline{\varepsilon}]$. For all such ε, the conditional expectation (20) is strictly positive so this change in the approval set would strictly increase expected consumer surplus. A similar argument applies if $\underline{\Delta CS}_{k'} < E^{\mathcal{A}}_{k'}(\overline{a}_{k'})$.

Step 4. Next, we argue that for all $j < k$ such that $j, k \in \mathcal{K}^+$ it must be that $\underline{\Delta \Pi}_j \leq \underline{\Delta \Pi}_k$; that is, the bilateral profit change in the marginal merger by target j must be no greater than the bilateral profit change in the marginal merger by any larger target k. Figure 6(a) shows a situation that violates this condition, where the marginal merger by target 3 causes a smaller bilateral profit change, $\underline{\Delta \Pi}_3$, than the marginal merger by the smaller target 2, $\underline{\Delta \Pi}_2$.

For $j \in \mathcal{K}^+$, let $k' \equiv \arg\min_{k \in \mathcal{K}^+, k > j} \underline{\Delta \Pi}_k$ and suppose that $\underline{\Delta \Pi}_{k'} < \underline{\Delta \Pi}_j$. We know from the previous step that $\underline{\Delta CS}_{k'} = E^{\mathcal{A}}_{k'}(\overline{a}_{k'})$. Let $\overline{c}'_j$ be the post-merger cost level satisfying $\Delta \Pi(j, \overline{c}'_j) = \underline{\Delta \Pi}_{k'}$ and consider a change in the approval set from $\mathcal{A}$ to $\mathcal{A} \cup \overline{\mathcal{A}}^{\varepsilon}_j$ where

$$\overline{\mathcal{A}}^{\varepsilon}_j \equiv \{M_j : M_j = (j, \overline{c}_j) \text{ with } \overline{c}_j \in \left(\overline{c}'_j, \overline{c}'_j + \epsilon\right)\}.$$

The set $\overline{\mathcal{A}}^{\varepsilon}_j$ is shown in Figure 6(b). The change in expected consumer surplus from this change in the approval set equals $\Pr(M^*(\mathfrak{F},\mathcal{A} \cup \overline{\mathcal{A}}^{\varepsilon}_j) \in \overline{\mathcal{A}}^{\varepsilon}_j)$ times

$$E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F},\mathcal{A} \cup \overline{\mathcal{A}}^{\varepsilon}_j)) - E^{\mathcal{A}}_j(\overline{c}_j)|M^*(\mathfrak{F},\mathcal{A} \cup \overline{\mathcal{A}}^{\varepsilon}_j) \in \overline{\mathcal{A}}^{\varepsilon}_j], \tag{21}$$

where $\overline{c}_j$ is the realized cost level in the aggregate profit-maximizing merger $M^*(\mathfrak{F},\mathcal{A} \cup \overline{\mathcal{A}}^{\varepsilon}_j)$, which is a merger of firms 0 and j when the conditioning statement is satisfied. As $\varepsilon \to 0$, the expected change in (21) converges to

$$\begin{aligned} \Delta CS(j, \overline{c}'_j) - E^{\mathcal{A}}_j(\overline{c}'_j) &= \Delta CS(j, \overline{c}'_j) - E^{\mathcal{A}}_{k'}(\overline{a}_{k'}) \\ &> \underline{\Delta CS}_{k'} - E^{\mathcal{A}}_{k'}(\overline{a}_{k'}) \\ &= 0, \end{aligned}$$

where the inequality follows from Corollary 1 since $\Delta \Pi(j, \overline{c}'_j) = \underline{\Delta \Pi}_{k'}$.

Step 5. We next argue that $\underline{\Delta CS}_j < \underline{\Delta CS}_k$ for all $j, k \in \mathcal{K}^+$ with $j < k$. Suppose otherwise; i.e., for some $j, h \in \mathcal{K}^+$ with $h > j$ we have $\underline{\Delta CS}_j \geq \underline{\Delta CS}_h$. Define $k = \arg\min \left\{h \in \mathcal{K}^+ : h > j \text{ and } \underline{\Delta CS}_j \geq \underline{\Delta CS}_h\right\}$. Figure 7 depicts such a situation where $j = 2$ and $k = 3$.

By Step 3, we must have $E^{\mathcal{A}}_j(\overline{a}_j) = \underline{\Delta CS}_j \geq \underline{\Delta CS}_k = E^{\mathcal{A}}_k(\overline{a}_k)$. But recalling (17), $E^{\mathcal{A}}_k(\overline{a}_k)$ can be written as a weighted average of two conditional expectations:

$$E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F} \backslash M_k,\mathcal{A}))|M_k = (k, \overline{a}_k),\ M_k = M^*(\mathfrak{F},\mathcal{A}), \text{ and } \Delta \Pi(M^*(\mathfrak{F} \backslash M_k,\mathcal{A})) < \underline{\Delta \Pi}_j] \tag{22}$$

and

$$E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F} \backslash M_k,\mathcal{A}))|M_k = (k, \overline{a}_k),\ M_k = M^*(\mathfrak{F},\mathcal{A}), \text{ and } \Delta \Pi(M^*(\mathfrak{F} \backslash M_k,\mathcal{A})) \in [\underline{\Delta \Pi}_j, \underline{\Delta \Pi}_k]]. \tag{23}$$

Expectation (22) conditions on the event that the next-most-profitable merger other than $(k, \overline{a}_k)$ induces a bilateral profit change less than $\underline{\Delta\Pi}_j$, the bilateral profit change of merger $(j, \overline{a}_j)$. Since no merger in $\mathcal{A}$ by either target k or j can have such a profit level (since $\underline{\Delta\Pi}_k \geq \underline{\Delta\Pi}_j$ by Step 4), the expectation (22) must exactly equal $E_j^{\mathcal{A}}(\overline{a}_j)$. Now consider the expectation (23). If $\Delta\Pi(M^*(\mathfrak{F}\backslash M_k,\mathcal{A})) \in [\underline{\Delta\Pi}_j, \underline{\Delta\Pi}_k)$, it could be that (i) $M^*(\mathfrak{F}\backslash M_k,\mathcal{A}) = (j, \overline{c}_j)$ for some $\overline{c}_j \leq \overline{a}_j$, or (ii) $M^*(\mathfrak{F}\backslash M_k,\mathcal{A}) = (r, \overline{c}_r)$ for some $r < j$, or (iii) $M^*(\mathfrak{F}\backslash M_k,\mathcal{A}) = (r, \overline{c}_r)$ for some $r > j$ and $r < k$. Now, in case (i) it is immediate that $\Delta CS(M^*(\mathfrak{F}\backslash M_k,\mathcal{A}) \geq \underline{CS}_j$, with strict inequality whenever $\overline{c}_j = \overline{a}_j$. In case (ii), the fact that $\Delta\Pi(r, \overline{c}_r) \geq \underline{\Delta\Pi}_j$ implies by Corollary 1 that

$$\Delta CS(M^*(\mathfrak{F}\backslash M_k,\mathcal{A}) = \Delta CS(r, \overline{c}_r) > \underline{CS}_j = E_j^{\mathcal{A}}(\overline{a}_j). \tag{24}$$

In case (iii), (24) follows from the definition of k. Thus, expectation (23) must strictly exceed $E_j^{\mathcal{A}}(\overline{a}_j)$, which leads to a contradiction.

Step 6. Finally, we argue that $\mathcal{K}^+ = \{1, ..., \widehat{K}\}$ for some $\widehat{K} \leq K$. To establish this fact, we show that if $k \notin \mathcal{K}^+$ and $k < K$, then $k+1 \notin \mathcal{K}^+$. We first observe that $\Delta CS(k,l) > \Delta CS(k+1,l)$, which follows because the profile of firms' costs following merger (k,l) are lower than following merger $(k+1,l)$ (the post-merger industry cost profile differs only for firms k and $k+1$, which have costs of l and c_{k+1} with the first merger and c_k and l with the second). Thus, if $k+1 \in \mathcal{K}^+$, then $\Delta CS(k+1, \overline{a}_{k+1}) < \Delta CS(k,l)$. But, an argument like that in Step 5 [using the fact that, by an argument like that in Step 3, $\Delta CS(k,l) \leq E_k^{\mathcal{A}}(\overline{a}_k)$] shows that $\Delta CS(k,l) < E_{k+1}^{\mathcal{A}}(\overline{a}_{k+1})$, so that $\Delta CS(k+1, \overline{a}_{k+1}) < E_{k+1}^{\mathcal{A}}(\overline{a}_{k+1})$, contradicting the conclusion of Step 3. $\square$

Proof of Proposition 2. Denote by $\mathcal{A}^*(\overline{\Delta\Pi}|J)$ a policy that is an element of $\arg\max_{\mathcal{A}\subseteq\Pi_{k\in J}[l,h_k]} ECS(\overline{\Delta\Pi}; \mathcal{A}, J)$ for a given J and $\overline{\Delta\Pi}$. Also, define $P(\overline{\Delta\Pi}|J,\mathcal{A}) \equiv \{k \in J : \Delta\Pi(k, \overline{a}_k) < \overline{\Delta\Pi}\}$ as the set of targets in J who may have an acceptable merger with profit below $\overline{\Delta\Pi}$ under policy $\mathcal{A} \subseteq \Pi_{k\in J}[l, h_k]$. Note that changes to $\mathcal{A}$ that alter acceptance sets only for $k \notin P(\overline{\Delta\Pi}|J,\mathcal{A})$ and leave $P(\overline{\Delta\Pi}|J,\mathcal{A})$ unchanged have no effect on the value of $ECS(\overline{\Delta\Pi}; \mathcal{A}, J)$. Finally, for any set $\mathcal{A}$, let $\mathcal{A}_J \equiv \Pi_{j\in J}\mathcal{A}_j$.

With these preliminaries, we now establish the result. Observe first that, for any J, a sufficient condition for M_k with $k \in J$ and $\Delta\Pi(M_k) < \overline{\Delta\Pi}$ to be approved in *any* solution to $\max_{\mathcal{A}\subseteq\Pi_{j\in J}[l,h_j]} ECS(\overline{\Delta\Pi}; \mathcal{A}, J)$ is that its CS-level, $\Delta CS(M_k)$, strictly exceeds $\max_{\mathcal{A}\subseteq\Pi_{j\in J\backslash k}[l,h_j]} ECS(\Delta\Pi(M_k); \mathcal{A}, J\backslash k)$. We will establish the result through an induction argument that shows that for all $\overline{\Delta\Pi}$ and any J such that $1 \in J$, if $\mathcal{A}^*(\overline{\Delta\Pi}|J) \in \arg\max_{\mathcal{A}\subseteq\Pi_{j\in J}[l,h_j]} ECS(\overline{\Delta\Pi}; \mathcal{A}, J)$ then

$$P(\overline{\Delta\Pi}|J, \mathcal{A}^*(\overline{\Delta\Pi}|J)) = P(\overline{\Delta\Pi}|J, \mathcal{A}^C(J)) \tag{25}$$

and

$$\mathcal{A}_k^*(\overline{\Delta\Pi}|J) = \mathcal{A}_k^C(J) \text{ for all } k \in P(\overline{\Delta\Pi}|J, \mathcal{A}^C(J)). \tag{26}$$

That is, any policy $\mathcal{A}$ that maximizes $ECS(\overline{\Delta\Pi}; \mathcal{A}, J)$ accepts with positive probability [conditional on the most profitable acceptable merger having $\Delta\Pi(M_j) \leq \overline{\Delta\Pi}$] mergers involving the same set of targets as does the cutoff policy $\mathcal{A}^C(J)$, and coincides with the cutoff policy $\mathcal{A}^C(J)$ for all such targets. In particular, this implies that the cutoff policy $\mathcal{A}^C(J) \in \arg\max_{\mathcal{A}\subseteq\Pi_{k\in J}[l,h_k]} ECS(\overline{\Delta\Pi}, \mathcal{A}, J)$ for all $\overline{\Delta\Pi}$ and any J such that $1 \in J$. Taking $\overline{\Delta\Pi} = \infty$ and $J = \mathcal{K}$ will then yield the result.

Consider first the set $J = \{1\}$. Then, we have $\overline{a}_1^C(J) = \widehat{c}_1(Q^0)$. Moreover, it is immediate – given our earlier discussion – that (25) and (26) hold for all $\overline{\Delta\Pi}$.

Now consider any set $J = J_n$ with $\#J_n = n$ and $1 \in J_n$, and assume:

Induction Hypothesis 1: Properties (25) and (26) hold for any set $J = J_{n'}$ with $1 \in J_{n'}$ and $n' < n$.

Number the targets in set J_n in increasing order of their pre-merger market share as $(1, ..., n)$. If $P(\overline{\Delta\Pi}|J, \mathcal{A}^C(J)) = \varnothing$, then $\overline{\Delta\Pi} \leq \Delta\Pi(1, \widehat{c}_1(Q_0))$. From Proposition 1, it follows immediately that $P(\overline{\Delta\Pi}|J, \mathcal{A}^*(\overline{\Delta\Pi}|J)) = \varnothing$. Hence, properties (25) and (26) hold for set J_n.

So suppose now instead that $P(\overline{\Delta\Pi}|J, \mathcal{A}^C(J)) \neq \varnothing$. Note that, since $\Delta\Pi(k, \overline{a}_k^C(J))$ is increasing in k, the set $P(\overline{\Delta\Pi}|J, \mathcal{A}^C(J))$ is of the form $P(\overline{\Delta\Pi}|J, \mathcal{A}^C(J)) = \{1, ..., \widehat{j}(\overline{\Delta\Pi})\}$ for some $\widehat{j}(\overline{\Delta\Pi})$.

Consider first the treatment of mergers with target 1. We have $\overline{a}_1^C(J) = \widehat{c}_1(Q^0)$. Moreover, the following two properties hold for all $\overline{\Delta\Pi}$: for any $\mathcal{A}^*(\overline{\Delta\Pi}|J_n) \in \arg\max_{\mathcal{A} \subseteq \Pi_{j \in J_n}[l, h_j]} ECS(\overline{\Delta\Pi}; \mathcal{A}, J_n)$,

$$1 \in P(\overline{\Delta\Pi}|J_n, \mathcal{A}^*(\overline{\Delta\Pi}|J_n)) \Leftrightarrow 1 \in P(\overline{\Delta\Pi}|J_n, \mathcal{A}^C(J_n)) \tag{27}$$

and

$$\mathcal{A}_1^*(\overline{\Delta\Pi}|J_n) = \mathcal{A}_1^C(J_n) \text{ if } 1 \in P(\overline{\Delta\Pi}|J_n, \mathcal{A}^C(J_n)). \tag{28}$$

These follow from the following facts: (i) No CS-decreasing merger M_1 can be accepted in $\mathcal{A}^*(\overline{\Delta\Pi}|J_n)$; (ii) for any $\mathcal{A} \subseteq \Pi_{j \in J_n}[l, h_j]$, $ECS(\Delta\Pi(1, \overline{c}_1); \mathcal{A}, J_n) < \Delta CS(1, \overline{c}_1)$ for all $\overline{c}_1 < \widehat{c}_1(Q^0)$, so all mergers $M_1 = (1, \overline{c}_1)$ such that $\overline{c}_1 < \widehat{c}_1(Q^0)$ and $\Delta\Pi(1, \overline{c}_1) \leq \overline{\Delta\Pi}$ must be in $\mathcal{A}^*(\overline{\Delta\Pi}|J_n)$, and (iii) accepting all mergers M_1 such that $\Delta\Pi(1, \overline{c}_1) > \overline{\Delta\Pi}$ maximizes $\Pr(\Delta\Pi(M_1) > \overline{\Delta\Pi}$ and $M_1 \in \mathcal{A}_1)$ and, since accepting the mergers described in (ii) is optimal, therefore maximizes $ECS(\overline{\Delta\Pi}; \mathcal{A}, J_n)$.

Now, consider a merger with target $k > 1$ and assume:

Induction Hypothesis 2: For all $k' < k$, the following two properties hold for all $\overline{\Delta\Pi}$: for any $\mathcal{A}^*(\overline{\Delta\Pi}|J_n) \in \arg\max_{\mathcal{A} \subseteq \Pi_{j \in J_n}[l, h_j]} ECS(\overline{\Delta\Pi}; \mathcal{A}, J_n)$,

$$k' \in P(\overline{\Delta\Pi}|J_n, \mathcal{A}^*(\Delta\Pi|J_n)) \Leftrightarrow k' \in P(\overline{\Delta\Pi}|J_n, \mathcal{A}^C(J_n)) \tag{29}$$

and

$$\mathcal{A}_{k'}^*(\overline{\Delta\Pi}|J_n) = \mathcal{A}_{k'}^C(J_n) \text{ if } k' \in P(\overline{\Delta\Pi}|J_n, \mathcal{A}^C(J_n)). \tag{30}$$

We will show that properties (29) and (30) hold as well for k so that Induction Hypothesis 2 holds for $k + 1$. Suppose, first, that $k \notin P(\overline{\Delta\Pi}|J_n, A^C(J_n))$. Then every M_k with $\Delta\Pi(M_k) \leq \overline{\Delta\Pi}$ has $ECS(\Delta\Pi(M_k); \mathcal{A}_{J_n \setminus k}^C(J_n), J_n \setminus k) > \Delta CS(M_k)$. But by Induction Hypothesis 2 and Proposition 1 [which implies that in $\mathcal{A}^*(\overline{\Delta\Pi}|J_n)$ we must have $\underline{\Delta\Pi}_k < \underline{\Delta\Pi}_j$ for any $j > k$ such that $j \in P(\overline{\Delta\Pi}|J_n, \mathcal{A}^*(\overline{\Delta\Pi}|J_n))$], which implies that $ECS(\Delta\Pi(M_k); \mathcal{A}_{J_n \setminus k}^*(\overline{\Delta\Pi}|J_n), J_n \setminus k) = ECS(\Delta\Pi(M_k); \mathcal{A}_{J_n \setminus k}^C(J_n), J_n \setminus k)$. Hence, merger M_k cannot be in $\mathcal{A}^*(\overline{\Delta\Pi}|J_n)$; i.e., $k \notin P(\overline{\Delta\Pi}|J_n, \mathcal{A}^*(\overline{\Delta\Pi}|J_n))$.

Suppose now instead that $k \in P(\overline{\Delta\Pi}|J, A^C(J_n))$. Observe, first, that every $M_k = (k, \overline{c}_k)$ with $\overline{c}_k > \overline{a}_k^C(J_n)$ has $ECS(\Delta\Pi(M_k); \mathcal{A}_{J_n \setminus k}^C(J_n), J_n \setminus k) > \Delta CS(M_k)$, and since by Induction Hypothesis 2 and Proposition 1, $ECS(\Delta\Pi(M_k); \mathcal{A}_{J_n \setminus k}^*(\overline{\Delta\Pi}|J_n), J_n \setminus k) = ECS(\Delta\Pi(M_k); \mathcal{A}_{J_n \setminus k}^C(J_n), J_n \setminus k)$, the merger cannot be in $\mathcal{A}^*(\overline{\Delta\Pi}|J_n)$; i.e., $\mathcal{A}_k^*(\overline{\Delta\Pi}|J_n) \subseteq \mathcal{A}_k^C(J_n)$. Next, consider mergers $M_k = (k, \overline{c}_k)$ with $\overline{c}_k < \overline{a}_k^C(J_n)$. Condition (9) combined with Induction Hypotheses 1 and 2 imply that each of these mergers satisfies $\Delta CS(M_k) > ECS(\Delta\Pi(M_k); \mathcal{A}^C(J_n \setminus k), J_n \setminus k) = \max_{\mathcal{A} \subseteq \Pi_{j \in J_n \setminus k}[l, h_j]} ECS(\Delta\Pi(M_k), \mathcal{A}, J_n \setminus k)$, and hence must be included in $\mathcal{A}^*(\overline{\Delta\Pi}|J_n)$; i.e., $\mathcal{A}_k^C(J_n) \subseteq \mathcal{A}_k^*(\overline{\Delta\Pi}|J_n)$. We thus have $\mathcal{A}_k^C(J_n) = \mathcal{A}_k^*(\overline{\Delta\Pi}|J_n)$. Hence, properties (29) and (30) hold as well for k. Applying induction (twice) then yields the result. $\square$

Proof of Proposition 3. Let $\mathcal{A}$ denote the optimal approval policy with cutoffs $(\overline{a}_1, ..., \overline{a}_K)$ when $\Pr(\phi_k = 1) = \theta_k$, and let $\mathcal{A}'$ denote the optimal approval policy with cutoffs $(\overline{a}'_1, ..., \overline{a}'_K)$ when $\Pr(\phi_k = 1) = \theta'_k$. From the recursive definition of the cutoffs, it follows immediately that a change in θ_k does not affect the cutoffs for any smaller merger M_j, $j < k$, nor the cutoff of merger M_k itself. Hence, $\underline{\Delta CS}_j{}' = \underline{\Delta CS}_j$ for all $j \leq k$.

Consider now the cutoff for merger M_{k+1}, $k+1 \leq \hat{K}$. We can write the cutoff condition as

$$\begin{aligned}
\underline{\Delta CS}_{k+1} \quad = \quad & \Pr(\phi_k = 1 |\ \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1})) \\
& \times E_{\mathfrak{F}_{\{1,...,k\}}}\left[\Delta CS\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \mid \right. \\
& \left. \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{j+1}) \text{ and } \phi_k = 1\right] \\
& + \left[1 - \Pr(\phi_k = 1 |\ \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1})\right] \\
& \times E_{\mathfrak{F}_{\{1,...,k\}}}\left[\Delta CS\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \mid \right. \\
& \left. \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1}) \text{ and } \phi_k = 0\right],
\end{aligned}$$

where $\mathcal{A}_J \equiv \Pi_{j\in J}\mathcal{A}_j$.

Note first that the optimal policy must be such that

$$\begin{aligned}
& E_{\mathfrak{F}_{\{1,...,k\}}}\left[\Delta CS\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \mid M_{k+1} = (k+1, \overline{a}_{k+1}), \right. \\
& \left. \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(M_{k+1}), \text{ and } \phi_k = 1\right] \\
> \quad & E_{\mathfrak{F}_{\{1,...,k\}}}\left[\Delta CS\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \mid M_{k+1} = (k+1, \overline{a}_{k+1}), \right. \\
& \left. \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(M_{k+1}), \text{ and } \phi_k = 0\right].
\end{aligned}$$

To see this, consider the case where $\phi_k = 1$ and $\Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1})$. Two cases can arise: (i) $M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right) \neq M_k$ and (ii) $M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right) = M_k$. In case (i) the outcome is the same as when M_k were not feasible ($\phi_k = 0$). In case (ii), merger M_k will be implemented. If merger M_k were not feasible, we would instead obtain the expected consumer surplus of the next most profitable allowable merger. By the optimality of the approval policy, $\Delta CS(M_k)$ must weakly exceed (and, generically, strictly) the expected consumer surplus of the next-most profitable allowable merger.

Next, note that we can rewrite the conditional probability as

$$\begin{aligned}
& \Pr(\phi_k = 1 | \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1})) \\
= \quad & \Pr(\ \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1}) | \phi_k = 1)\theta_k \\
& \times \left\{\Pr(\Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1}) | \phi_k = 1)\theta_k \right. \\
& \left. + \Pr(\Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1}) | \phi_k = 0)(1-\theta_k)\right\}^{-1} \\
= \quad & \left\{1 + \frac{\Pr(\Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1}) | \phi_k = 0)}{\Pr(\Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1}) | \phi_k = 1)}\left(\frac{1-\theta_k}{\theta_k}\right)\right\}^{-1}.
\end{aligned}$$

Hence, an increase in θ_k induces an increase in the conditional probability $\Pr(\phi_k = 1 | \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,...,k\}}, \mathcal{A}_{\{1,...,k\}}\right)\right) \leq \Delta\Pi(k+1, \overline{a}_{k+1}))$. But this implies that an increase in θ_k induces an increase in the RHS of the cutoff condition for merger M_{k+1}. Hence, $\underline{\Delta CS}'_{k+1} > \underline{\Delta CS}_{k+1}$.

Consider now the induction hypothesis that $\underline{\Delta CS}'_{k'} > \underline{\Delta CS}_{k'}$ for all $k < k' < j \leq \hat{K}$. In particular, $\underline{\Delta CS}'_{j-1} > \underline{\Delta CS}_{j-1}$. We claim that this implies that $\underline{\Delta CS}'_j > \underline{\Delta CS}_j$. To see this, note that we can decompose the effect of the increase in θ_k on the conditional expectation of the next-most profitable merger into two steps:

1. Increase the feasibility probability from θ_k to $\theta_k' > \theta_k$, holding fixed the approval policy $\mathcal{A}$.

2. Change the approval policy from $\mathcal{A}$ to $\mathcal{A}'$.

Consider first step (1). For the same reason as before, the increase in the feasibility probability must raise the conditional expectation

$$E_{\mathfrak{F}_{\{1,\ldots,j-1\}}}\left[\Delta CS\left(M^*\left(\mathfrak{F}_{\{1,\ldots,j-1\}},\mathcal{A}_{\{1,\ldots,j-1\}}\right)\right) \mid \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,\ldots,j-1\}},\mathcal{A}_{\{1,\ldots,j-1\}}\right)\right) \leq \Delta\Pi(j,\overline{a}_j)\right]$$

by the optimality of the approval policy $\mathcal{A}$.

Consider now step (2). The outcome under the two policies differs only in the event where $M^*\left(\mathfrak{F}_{\{1,\ldots,j-1\}},\mathcal{A}_{\{1,\ldots,j-1\}}\right) \notin \mathcal{A}'$. Let $M_i = M^*\left(\mathfrak{F}_{\{1,\ldots,j-1\}},\mathcal{A}_{\{1,\ldots,j-1\}}\right)$. Under policy $\mathcal{A}$, the outcome in this event is $\Delta CS(M_i)$. Under policy $\mathcal{A}'$ instead, the expected outcome is

$$E_{\mathfrak{F}_{\{1,\ldots,i-1\}}}\left[\Delta CS\left(M^*\left(\mathfrak{F}_{\{1,\ldots,i-1\}},\mathcal{A}'_{\{1,\ldots,i-1\}}\right)\right) \mid \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,\ldots,i-1\}},\mathcal{A}'_{\{1,\ldots,i-1\}}\right)\right) \leq \Delta\Pi(k,\overline{c}_i)\right].$$

But as $M_i \notin \mathcal{A}'$, we must have

$$\begin{aligned} & E_{\mathfrak{F}_{\{1,\ldots,i-1\}}}\left[\Delta CS\left(M^*\left(\mathfrak{F}_{\{1,\ldots,i-1\}},\mathcal{A}'_{\{1,\ldots,i-1\}}\right)\right) \mid \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1,\ldots,i-1\}},\mathcal{A}'_{\{1,\ldots,i-1\}}\right)\right) \leq \Delta\Pi(j,\overline{c}_i)\right] \\ > \; & \Delta CS(M_i). \end{aligned}$$

As the expected consumer surplus increases at each step, we must have $\underline{\Delta CS}'_j > \underline{\Delta CS}_j$. □

Proof of Proposition 4. A change in firm 0's marginal cost does not affect the outcome (consumer surplus, profits) after any merger M_k, $k \geq 1$, but it does affect the pre-merger outcome. In particular, we have $Q^{0\prime} > Q^0$ so that $\gamma \equiv CS^{0\prime} - CS^0 > 0$. Let $\eta_k \equiv [\pi_0^{0\prime} + \pi_k^{0\prime}] - [\pi_0^0 + \pi_k^0]$ denote the induced change in the joint pre-merger profit of firms 0 and k. The key observation is that the profit of a more efficient firm falls by a larger amount than that of a less efficient as price falls. That is, η_k is decreasing in k.

Consider first merger M_1. We have $\Delta CS(1,\overline{a}_1')' = \Delta CS(1,\overline{a}_1') - \gamma = 0$. Hence, $\Delta CS(1,\overline{a}_1') > \Delta CS(1,\overline{a}_1) = 0$, implying that $\overline{a}_1' < \overline{a}_1$. Consider now the (marginal) merger $M_2 = (2,\overline{a}_2)$. Let $(1,\widetilde{a}_1)$ be such that $\Delta\Pi(1,\widetilde{a}_1) = \Delta\Pi(2,\overline{a}_2)$, and $(1,\widetilde{a}_1')$ be such that $\Delta\Pi(1,\widetilde{a}_1')' = \Delta\Pi(2,\overline{a}_2)'$. We have

$$\begin{aligned} \Delta\Pi(1,\widetilde{a}_1)' &= \Delta\Pi(1,\widetilde{a}_1) - \eta_1 \\ &< \Delta\Pi(1,\widetilde{a}_1) - \eta_2 \\ &= \Delta\Pi(2,\overline{a}_2) - \eta_2 \\ &= \Delta\Pi(2,\overline{a}_2)' \\ &= \Delta\Pi(1,\widetilde{a}_1')', \end{aligned}$$

where the inequality follows from $\eta_1 > \eta_2$. Hence, $\widetilde{a}_1' < \widetilde{a}_1$. That is, before the reduction in c_0, any merger M_1 with $\overline{c}_1 \geq \widetilde{a}_1$ induced a smaller increase in bilateral profit than merger $M_2 = (2,\overline{a}_2)$. After the reduction in c_0, this is still true, but now – in addition – any merger M_1 with $\widetilde{a}_1 > \overline{c}_1 \geq \widetilde{a}_1'$ also induces a smaller increase in bilateral profit than merger $M_2 = (2,\overline{a}_2)$. That is, there are now more and (in an FOSD sense) more efficient mergers M_1 that are less profitable than $M_2 = (2, \overline{a}_2)$. Since the induced CS-increase of merger M_1 is the greater, the lower is $\overline{c}_1$, we thus have again that

$$
\begin{aligned}
& E_{\mathfrak{F}_{\{1\}}}\left[\Delta CS\left(M^*\left(\mathfrak{F}_{\{1\}}, \mathcal{A}'_{\{1\}}\right)\right)' \mid \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1\}}, \mathcal{A}'_{\{1\}}\right)'\right) \leq \Delta\Pi(2, \overline{a}_2)'\right] \\
> \; & E_{\mathfrak{F}_{\{1\}}}\left[\Delta CS\left(M^*\left(\mathfrak{F}_{\{1\}}, \mathcal{A}_{\{1\}}\right)\right) \mid \Delta\Pi\left(M^*\left(\mathfrak{F}_{\{1\}}, \mathcal{A}_{\{1\}}\right)\right) \leq \Delta\Pi(2, \overline{a}_2)\right] - \gamma \\
= \; & \Delta CS(2, \overline{a}_2) - \gamma \\
= \; & \Delta CS(2, \overline{a}_2)'.
\end{aligned}
$$

Hence, $\overline{a}'_2 < \overline{a}_2$. Under the induction hypothesis that $\overline{a}'_j < \overline{a}_j$ for every $j < k \leq \hat{K}$, a similar argument can be used to show that $\overline{a}'_k < \overline{a}_k$. □

Lemma 7. *Consider the function* $H(s_1, ..., s_N) = \sum_n (s_n)^2$ *and two vectors* $s' = (s'_1, ..., s'_N)$ *and* $s'' = (s''_1, ..., s''_N)$ *having* $\sum_{n=1}^N s'_n = \sum_{n=1}^N s''_n$. *If for some* r, *(i)* $s'_r \geq s'_j$ *for all* $j \neq r$, *(ii)* $s''_r > s'_r$, *and (iii)* $s''_j \leq s'_j$ *for all* $j \neq r$, *then* $H(s'') > H(s')$.

Proof. Without loss of generality, take $r = 1$ and define $\Delta_n \equiv s'_n - s''_n$ for $n > 1$. Observe that $\Delta_n \geq 0$ for all $n > 1$ and $\Delta_n > 0$ for some $n > 1$. Define as well the vectors $s^n \equiv (s'_1 + \sum_{t=2}^n \Delta_t, s'_2 - \Delta_2, ..., s'_n - \Delta_n, s'_{n+1}, ..., s'_N)$ for $n > 1$ and $s^1 \equiv s'$. Note that $s^N = s''$. Then

$$H(s'') - H(s') = \sum_{n=1}^{N-1} [H(s^{n+1}) - H(s^n)].$$

Now letting $\overline{s}_1^1 \equiv s'_1$ and $\overline{s}_1^n \equiv s'_1 + \sum_{t=2}^n \Delta_t \geq s'_1$ for all $n > 1$, each term in this sum is nonnegative,

$$
\begin{aligned}
H(s^{n+1}) - H(s^n) &= (\overline{s}_1^n + \Delta_{n+1})^2 + (s'_{n+1} - \Delta_{n+1})^2 - (\overline{s}_1^n)^2 - (s'_{n+1})^2 \\
&= 2\Delta_{n+1}(\overline{s}_1^n - s'_{n+1}) + 2(\Delta_{n+1})^2 \geq 0,
\end{aligned}
$$

and strictly positive if $\Delta_{n+1} > 0$. Since $\Delta_{n+1} > 0$ for some $n \geq 1$, the result follows. □

Proof of Lemma 4. From the discussion in the main text, the post-merger aggregate profit is given by (10). As both mergers induce the same level of consumer surplus (and thus the same Q), the first term on the right-hand side of (10) is the same for both mergers. It thus suffices to show that the larger merger M_k induces a larger value of H than the smaller merger M_j.

Now, as both mergers induce the same Q, Assumption 1 implies that the output of any firm not involved in M_j or M_k is the same under both mergers. Hence,

$$s_k(M_k) + s_j(M_k) = s_k(M_j) + s_j(M_j). \tag{31}$$

Next, recall that a CS-nondecreasing merger increases the share of the merging firms and reduces the share of all nonmerging firms. Thus, we have $s_k(M_k) \geq s_k + s_0 > s_k(M_j)$ and $s_j(M_j) \geq s_j + s_0 > s_j(M_k)$. In addition, since total output is the same after both mergers and $c_k < c_j$, we also have $s_j(M_k) < s_k(M_j)$. By (31), this in turn implies that $s_k(M_k) > s_j(M_j)$. Hence, the distribution of market shares after the larger merger M_k is a sum-preserving spread of those after the smaller merger M_j:

$$s_k(M_k) > \max\{s_j(M_j), s_k(M_j)\} \geq \min\{s_j(M_j), s_k(M_j)\} > s_j(M_k). \tag{32}$$

By Lemma 7 in the Appendix (just above), H is therefore larger after M_k than after M_j. □

Lemma 8. *The slope of the curve for merger M_k in $(\Delta\Pi, \Delta CS)$-space is given by*

$$\frac{d\Delta\Pi(M_k)}{d\Delta CS(M_k)} = -2 - \left[\frac{P''(Q(M_k))Q(M_k)}{P'(Q(M_k))}\right] H(M_k) + \left[\frac{2}{P'(Q(M_k))Q(M_k)}\right]\left[\frac{r(Q(M_k);\overline{c}_{M_k})}{dQ(M_k)/d\overline{c}_{M_k}}\right],$$

where $r(Q;c) \equiv \{q|P(Q) - c + qP'(Q) = 0\}$ is the "cumulative best reply" of a firm with marginal cost c to aggregate output Q.

Proof. The change in ΔCS induced by a small increase in post-merger marginal cost is

$$\frac{d\Delta CS(M_k)}{d\overline{c}_{M_k}} = -P'(\overline{Q})\overline{Q}\frac{d\overline{Q}}{d\overline{c}_{M_k}},$$

where $\overline{Q} \equiv Q(M_k)$ is aggregate output following merger M_k. Recall that aggregate profit can be written as $\eta(\overline{Q})\overline{H}$, where $\overline{H} \equiv H(M_k)$ is the post-merger Herfindahl index and $\eta(\overline{Q}) \equiv -P'(\overline{Q})\overline{Q}^2$. The effect of a small increase in post-merger marginal cost on the change in aggregate profit induced by merger M_k is thus given by

$$\frac{d\Delta\Pi(M_k)}{d\overline{c}_{Mk}} = \eta'(\overline{Q})\frac{d\overline{Q}}{d\overline{c}_{M_k}}\overline{H} + \eta(\overline{Q})\frac{d\overline{H}}{d\overline{c}_{M_k}},$$

where

$$\eta'(\overline{Q}) = -[P''(\overline{Q})\overline{Q}^2 + 2P'(\overline{Q})\overline{Q}].$$

Putting this together, we obtain

$$\begin{aligned}\frac{d\Delta\Pi(M_k)}{d\Delta CS(M_k)} &= -\left[\frac{\eta'(\overline{Q})}{P'(\overline{Q})\overline{Q}}\right]\overline{H} - \left[\frac{\eta(\overline{Q})}{P'(\overline{Q})\overline{Q}}\right]\frac{(d\overline{H}/d\overline{c}_{M_k})}{(d\overline{Q}/d\overline{c}_{M_k})} \\ &= \left[2 + \frac{P''(\overline{Q})\overline{Q}}{P'(\overline{Q})}\right]\overline{H} + \overline{Q}\frac{(d\overline{H}/d\overline{c}_{M_k})}{(d\overline{Q}/d\overline{c}_{M_k})} \qquad (33)\end{aligned}$$

Now, we have

$$\begin{aligned}\frac{d\overline{H}}{d\overline{c}_{M_k}} &= \frac{d}{d\overline{c}_{M_k}}\left[\frac{\sum_i r(\overline{Q};c_i)^2}{\overline{Q}^2}\right] \\ &= -\left(\frac{2}{\overline{Q}^3}\right)\frac{d\overline{Q}}{d\overline{c}_{M_k}}\left[\sum_i r(\overline{Q};c_i)^2\right] + \left(\frac{1}{\overline{Q}^2}\right)\left\{2r(\overline{Q};\overline{c}_{M_k})\frac{\partial r(\overline{Q};\overline{c}_{M_k})}{\partial\overline{c}_{M_k}} + 2\sum_i r(\overline{Q};c_i)\frac{dr(\overline{Q};c_i)}{d\overline{Q}}\frac{d\overline{Q}}{d\overline{c}_{M_k}}\right\} \\ &= -\left(\frac{2}{\overline{Q}}\right)\overline{H}\frac{d\overline{Q}}{d\overline{c}_{M_k}} + \left(\frac{2}{P'(\overline{Q})\overline{Q}^2}\right)r(\overline{Q};\overline{c}_{M_k}) - \left(\frac{2}{\overline{Q}^2}\right)\frac{d\overline{Q}}{d\overline{c}_{M_k}}\sum_i\left[r(\overline{Q};c_i) + \frac{P''(\overline{Q})}{P'(\overline{Q})}r(\overline{Q};c_i)^2\right] \\ &= -\left(\frac{2}{\overline{Q}}\right)\overline{H}\frac{d\overline{Q}}{d\overline{c}_{M_k}} + \left(\frac{2}{P'(\overline{Q})\overline{Q}^2}\right)r(\overline{Q};\overline{c}_{M_k}) - \left(\frac{2}{\overline{Q}}\right)\frac{d\overline{Q}}{d\overline{c}_{M_k}} - 2\frac{P''(\overline{Q})}{P'(\overline{Q})}\overline{H}\frac{d\overline{Q}}{d\overline{c}_{M_k}}, \qquad (34)\end{aligned}$$

where the third equality follows using the facts that $\partial r(\overline{Q};\overline{c}_{M_k})/\partial\overline{c}_{M_k} = 1/P'(\overline{Q})$ and $dr(\overline{Q};c_i)/d\overline{Q} = -\left(1 + r(\overline{Q};c_i)P''(\overline{Q})/P'(\overline{Q})\right)$. Thus, we have:

$$\overline{Q}\frac{(d\overline{H}/d\overline{c}_{M_k})}{(d\overline{Q}/d\overline{c}_{M_k})} = -2\overline{H} - 2 - 2\left(\frac{P''(\overline{Q})\overline{Q}}{P'(\overline{Q})}\right)\overline{H} + \left(\frac{2}{P'(\overline{Q})\overline{Q}}\right)\frac{r(\overline{Q};\overline{c}_{M_k})}{(d\overline{Q}/d\overline{c}_{M_k})}. \qquad (35)$$

Substituting (35) into (33), we obtain equation (11). □

Proof of Lemma 5. Let $\overline{Q} \equiv Q(M_k)$ denote post-merger aggregate output. Inserting

$$\frac{d\overline{Q}}{d\overline{c}_{M_k}} = \frac{1}{(N - m_k + 3)P'(\overline{Q}) + \overline{Q}P''(\overline{Q})}$$

into equation (11), we obtain

$$\begin{aligned} \frac{d\Delta\Pi(M_k)}{d\Delta CS(M_k)} &= -2 - \frac{\overline{Q}P''(\overline{Q})}{P'(\overline{Q})}\overline{H} + \frac{2\overline{s}_{M_k}}{P'(\overline{Q})}\left[(N - m_k + 3)P'(\overline{Q}) + \overline{Q}P''(\overline{Q})\right] \\ &= 2\left[(N - m_k + 3)\overline{s}_{M_k} - 1\right] + \frac{\overline{Q}P''(\overline{Q})}{P'(\overline{Q})}\left[2\overline{s}_{M_k} - \overline{H}\right], \end{aligned}$$

where $\overline{s}_{M_k}$ is the actual market share of the merged firm and $\overline{H} \equiv H(M_k)$ the actual post-merger Herfindahl index.

Now, we claim that $s_{M_k}^{naive} \geq H^{naive}(M_k)/2$ implies that $2\overline{s}_{M_k} \geq \overline{H}$. To see this, note that the naively-computed inequality $s_{M_k}^{naive} \geq H^{naive}(M_k)/2$ corresponds to the case of a CS-neutral merger. As merger M_k is CS-nondecreasing by assumption, it involves a (weakly) lower level of $\overline{c}_{M_k}$ (and a weakly greater level of aggregate output) than a CS-neutral merger. It therefore suffices to show that a small reduction in $\overline{c}_{M_k}$ leads to a larger value of $\left[2\overline{s}_{M_k} - \overline{H}\right]$, i.e., $d\left[2\overline{s}_{M_k} - \overline{H}\right] > 0$. But we have

$$\begin{aligned} d\left[2\overline{s}_{M_k} - \overline{H}\right] &= 2d\overline{s}_{M_k} - 2\left(\overline{s}_{M_k}d\overline{s}_{M_k} + \sum_{i \notin M_k}\overline{s}_i d\overline{s}_i\right) \\ &= 2(1 - \overline{s}_{M_k})\left(1 - \sum_{i \notin M_k} d\overline{s}_i\right) - 2\sum_{i \notin M_k}\overline{s}_i d\overline{s}_i \\ &> 0, \end{aligned}$$

where the inequality follows from the observation that the induced increase in aggregate output reduces the market share of each nonmerging firm i, i.e., $d\overline{s}_i < 0$.

Since Assumption 1 implies that $\overline{Q}P''(\overline{Q})/P'(\overline{Q}) > -1$, we obtain

$$\begin{aligned} \frac{d\Delta\Pi(M_k)}{d\Delta CS(M_k)} &\geq 2\left[(N - m_k + 3)\overline{s}_{M_k} - 1\right] - \left[2\overline{s}_{M_k} - \overline{H}\right] \\ &= 2\left[(N - m_k + 2)\overline{s}_{M_k} - 1\right] + \overline{H}. \end{aligned}$$

The r.h.s. of the last equation is positive if and only if

$$\frac{\overline{H}}{2} \geq 1 - (N - m_k + 2)\overline{s}_{M_k}.$$

We claim that this inequality is implied by the naively-computed analog,

$$\frac{H^{naive}(M_k)}{2} \geq 1 - (N - m_k + 2)s_{M_k}^{naive}.$$

To see this, consider the effect of decreasing the post-merger marginal cost $\overline{c}_{M_k}$ on $\left[\overline{H} - 2\left(1 - (N - m_k + 2)\overline{s}_{M_k}\right)\right]$:

$$\begin{aligned} d\left[\overline{H} - 2\left(1 - (N - m_k + 2)\overline{s}_{M_k}\right)\right] &= 2\left(\overline{s}_{M_k} d\overline{s}_{M_k} + \sum_{i \notin M_k} \overline{s}_i d\overline{s}_i\right) + 2(N - m_k + 2) d\overline{s}_{M_k} \\ &= 2(N - m_k + 2 + \overline{s}_{M_k})\left(1 - \sum_{i \notin M_k} d\overline{s}_i\right) + 2\sum_{i \notin M_k} \overline{s}_i d\overline{s}_i \\ &= 2\left\{(N - m_k + 2 + \overline{s}_{M_k}) - \sum_{i \notin M_k} (N - m_k + 2 + \overline{s}_{M_k} - \overline{s}_i) d\overline{s}_i\right\} \\ &> 0, \end{aligned}$$

where the inequality follows from the observation that $d\overline{s}_i < 0$ for all $i \notin M_k$ and $N - m_k + 2 \geq 1$. $\square$

Proof of Lemma 6. Let $q^0_{M_l} \equiv \sum_{i \in M_l} q^0_i$, $l = j, k$. The pre-merger first-order conditions imply that

$$[m_l P(Q^0) - \sum_{i \in M_l} c_i] + P'(Q^0) q^0_{M_l} = 0 \text{ for } l = j, k,$$

so

$$(m_k - m_j)P(Q^0) - \sum_{i \in M_k} c_i + \sum_{i \in M_j} c_i = P'(Q^0)(q^0_{M_j} - q^0_{M_k}) > 0. \tag{36}$$

Next, summing up the post-merger first-order conditions, we have

$$(N - m_l + 2)P(\overline{Q}) - \sum_i c_i + \sum_{i \in M_l} c_i - \overline{c}_{M_l} + P'(\overline{Q})\overline{Q} = 0 \text{ for } l = j, k, \tag{37}$$

where $N + 1$ is the number of firms prior to any merger. So,

$$-\left[(m_k - m_j)P(\overline{Q}) - \sum_{i \in M_k} c_i + \sum_{i \in M_j} c_i\right] - [\overline{c}_{M_k} - \overline{c}_{M_j}] = 0.$$

Since $-(m_k - m_j)P(\overline{Q}) \leq -(m_k - m_j)P(Q)$ as the mergers are CS-nondecreasing by assumption, we have

$$-\left[(m_k - m_j)P(Q) - \sum_{i \in M_k} c_i + \sum_{i \in M_j} c_i\right] - [\overline{c}_{M_k} - \overline{c}_{M_j}] \geq 0,$$

so (36) implies that $\overline{c}_{M_k} - \overline{c}_{M_j} < 0$, which in turn implies that $r(\overline{Q}; \overline{c}_{M_k}) > r(\overline{Q}; \overline{c}_{M_j})$.

Applying the implicit function theorem to (37), yields

$$\frac{d\overline{Q}}{d\overline{c}_{M_l}} = \frac{1}{(N - m_l + 3)P'(\overline{Q}) + \overline{Q}P''(\overline{Q})}.$$

As $m_k \leq m_j$, $P'(\overline{Q}) < 0$, $P'(\overline{Q}) + \overline{Q}P''(\overline{Q}) < 0$, and (from above) $r(\overline{Q}; \overline{c}_{M_k}) > r(\overline{Q}; \overline{c}_{M_j})$, we obtain

$$-\frac{r(\overline{Q}; \overline{c}_{M_k})}{d\overline{Q}/d\overline{c}_{M_k}} > -\frac{r(\overline{Q}; \overline{c}_{M_j})}{d\overline{Q}/d\overline{c}_{M_j}}.$$

Equation (11) implies that $d\Delta\Pi/d\Delta CS$ is larger for merger M_k than for M_j at any point where the curves cross, from which the assertion follows. $\square$

Proof of Lemma 4'. The proof proceeds exactly as that of Lemma 4, except that the inequalities $s_k(M_k) > s_k(M_j)$ and $s_j(M_j) > s_j(M_k)$ in equation (32) now hold because, by Assumption 3', any W-nondecreasing merger involves synergies ($\overline{c}_k < c_k$ and $\overline{c}_j < c_j$). □

Proof of Proposition 5. Steps 1-2 proceed along the same lines as those in the proof of Proposition 1.

Step 3. As in the absence of fixed cost savings, any optimal policy has the property that, for all $k \in \mathcal{K}^+$ and any $\overline{f}_k$, $\underline{\Delta CS}_k(\overline{f}_k)$ is equal to the expected change in consumer surplus from the next-most profitable merger $M^*(\mathfrak{F}\backslash(k, \overline{a}_k(\overline{f}_k), \overline{f}_k), \mathcal{A})$, conditional on the marginal merger $M_k = (k, \overline{a}_k(\overline{f}_k), \overline{f}_k)$ maximizing the change in the merging firms' bilateral profit in $\mathfrak{F}\cap\mathcal{A}$. That is,

$$\begin{aligned}
\underline{\Delta CS}_k(\overline{f}_k) &= E_k^{\mathcal{A}}(\overline{a}_k(\overline{f}_k), \overline{f}_k) \\
&\equiv E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F}\backslash M_k, \mathcal{A}))|M_k = (k, \overline{a}_k(\overline{f}_k), \overline{f}_k) \text{ and } \Delta\Pi(M^*(\mathfrak{F}\backslash M_k, \mathcal{A})) \leq \Delta\Pi(M_k)].
\end{aligned}$$

To see that this equation must hold for all $k \in \mathcal{K}^+$, suppose first that $\underline{\Delta CS}_{k'}(\overline{f}'_{k'}) > E_{k'}^{\mathcal{A}}(\overline{a}_{k'}(\overline{f}'_{k'}), \overline{f}'_{k'})$ for some firm $k' \in \mathcal{K}^+$ and fixed cost realization $\overline{f}'_{k'}$, and consider the alternative approval set $\mathcal{A} \cup \mathcal{A}_{k'}^{\varepsilon}$, where

$$\mathcal{A}_{k'}^{\varepsilon} \equiv \left\{ M_k : M_k = (k', \overline{c}_{k'}, \overline{f}_{k'}) \text{ with } \overline{c}_{k'} \in \left(\overline{a}_{k'}(\overline{f}'_{k'}), \overline{a}_{k'}(\overline{f}'_{k'}) + \varepsilon\right) \text{ and } \overline{f}_{k'} \in \left(\overline{f}'_{k'} - \varepsilon, \overline{f}'_{k'} + \varepsilon\right) \right\}.$$

Using the same type of argument as in the proof of Proposition 1, it is straightforward to show that, for $\varepsilon > 0$ small enough, the change in expected consumer surplus from changing the approval set from $\mathcal{A}$ to $\mathcal{A} \cup \mathcal{A}_{k'}^{\varepsilon}$ is strictly positive. A similar logic can be used to show that we cannot have $\underline{\Delta CS}_{k'}(\overline{f}'_{k'}) < E_{k'}^{\mathcal{A}}(\overline{a}_{k'}(\overline{f}'_{k'}), \overline{f}'_{k'})$.

Step 4. Let $\mathcal{M}_j^{CS} \equiv \{M_j : \Delta CS(M_j) = \underline{\Delta CS}_j \text{ and } M_j \in \mathcal{A}_j\}$ denote the set of marginal mergers M_j that induce a change in consumer surplus of $\underline{\Delta CS}_j$, and let $M_j^{CS} \in \mathcal{M}_j^{CS}$ denote the most profitable among these mergers, i.e., $\Delta\Pi(M_j^{CS}) \geq \Delta\Pi(M'_j)$ for all $M'_j \in \mathcal{M}_j^{CS}$. This merger is depicted in Figure 12 for $j = 2$. An optimal approval set must have the property that, for all $j < k$ such that $j, k \in \mathcal{K}^+$, we have $\Delta\Pi\left(M_j^{CS}\right) \leq \underline{\Delta\Pi}_k$. The argument is similar to (but slightly more involved than) Step 4 in the proof of Proposition 1: For $j \in \mathcal{K}^+$, let $k' \equiv \arg\min_{k\in\mathcal{K}^+, k>j} \underline{\Delta\Pi}_k$ and suppose that, contrary to our claim, $\underline{\Delta\Pi}_{k'} < \Delta\Pi(M_j^{CS})$. In Figure 12 we suppose that $k' = 3$. Let $M_{k'}^{\Pi} = (k', \overline{a}_{k'}(\overline{f}_{k'}^{\Pi}), \overline{f}_{k'}^{\Pi})$ denote the marginal merger $M_{k'}$ that induces the bilateral profit change $\underline{\Delta\Pi}_{k'}$, i.e., $\Delta\Pi\left(M_{k'}^{\Pi}\right) = \underline{\Delta\Pi}_{k'}$. By Step 3, $M_{k'}^{\Pi}$ is uniquely defined, and $\underline{\Delta CS}_{k'}(M_{k'}^{\Pi}) = E_{k'}^{\mathcal{A}}\left(\overline{a}_{k'}(\overline{f}_{k'}^{\Pi}), \overline{f}_{k'}^{\Pi}\right)$. Note that $E_{k'}^{\mathcal{A}}\left(\overline{a}_{k'}(\overline{f}_{k'}^{\Pi}), \overline{f}_{k'}^{\Pi}\right)$ can be written as a weighted average of

$$\tau_1 \equiv E_{\mathfrak{F}}[\Delta CS(M_j, \mathcal{A}))|M_{k'} = M_{k'}^{\Pi},\ M_j \in M^*(\mathfrak{F}\backslash M_{k'}, \mathcal{A}), \text{ and } \Delta\Pi(M^*(\mathfrak{F}\backslash M_{k'}, \mathcal{A})) \leq \Delta\Pi(M_{k'})]$$

and

$$\tau_2 \equiv E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F}\backslash M_{k'}, \mathcal{A}))|M_{k'} = M_{k'}^{\Pi},\ M_j \notin M^*(\mathfrak{F}\backslash M_{k'}, \mathcal{A}), \text{ and } \Delta\Pi(M^*(\mathfrak{F}\backslash M_{k'}, \mathcal{A})) \leq \Delta\Pi(M_{k'})],$$

where the probability weight on τ_1 is positive if and only if $\underline{\Delta\Pi}_j < \underline{\Delta\Pi}_{k'}$. Note also that $\tau_1 \geq \underline{\Delta CS}_j > \Delta CS\left(M_{k'}^{\Pi}\right) = E_{k'}^{\mathcal{A}}\left(\overline{a}_{k'}(\overline{f}_{k'}^{\Pi}), \overline{f}_{k'}^{\Pi}\right)$. Hence, by Step 3,

$$\Delta CS\left(M_{k'}^{\Pi}\right) = E_{k'}^{\mathcal{A}}\left(\overline{a}_{k'}(\overline{f}_{k'}^{\Pi}), \overline{f}_{k'}^{\Pi}\right) \geq \tau_2. \tag{38}$$

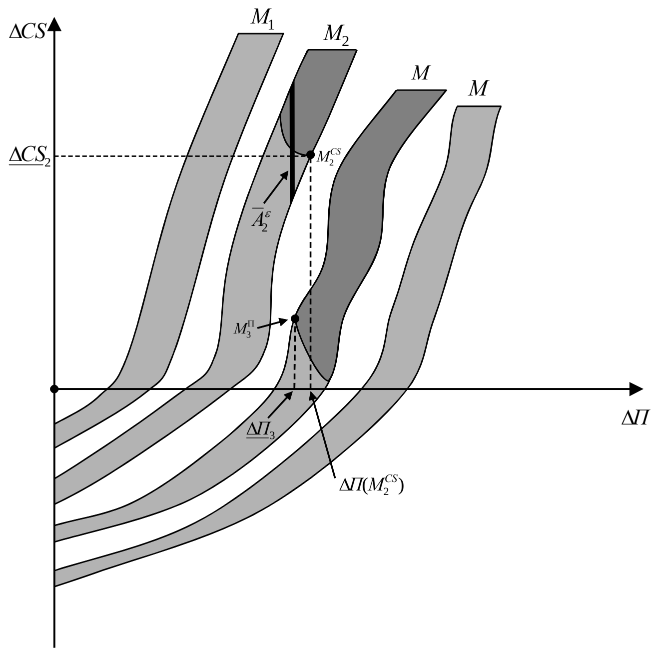

Figure 12: The figure illustrates the change considered in Step 4 of the proof of Proposition 5

Consider a change in the approval set from $\mathcal{A}$ to $\mathcal{A}\cup\overline{\mathcal{A}}_j^{\varepsilon}$, where

$$\overline{\mathcal{A}}_j^{\varepsilon} \equiv \{M_j : \Delta\Pi(M_j) \in [\underline{\Delta\Pi}_{k'} - \varepsilon, \underline{\Delta\Pi}_{k'}]\},$$

and $\varepsilon > 0$. Note that, as shown in Figure 12, $\overline{\mathcal{A}}_j^{\varepsilon} \not\subseteq \mathcal{A}$. The change in expected consumer surplus from this change in the approval set equals $\Pr(M^*(\mathfrak{F},\mathcal{A}\cup\overline{\mathcal{A}}_j^{\varepsilon}) \in (\mathcal{A}\cup\overline{\mathcal{A}}_j^{\varepsilon})\backslash\mathcal{A})$ [which is strictly positive as $\overline{\mathcal{A}}_j^{\varepsilon} \not\subseteq \mathcal{A}$] times

$$E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F},\mathcal{A}\cup\overline{\mathcal{A}}_j^{\varepsilon})) - E_j^{\mathcal{A}}(\overline{c}_j, \overline{f}_j) | M^*(\mathfrak{F},\mathcal{A}\cup\overline{\mathcal{A}}_j^{\varepsilon}) \in (\mathcal{A}\cup\overline{\mathcal{A}}_j^{\varepsilon})\backslash\mathcal{A}], \tag{39}$$

where $(\overline{c}_j, \overline{f}_j)$ is the pair of realized cost levels in the most profitable merger $M^*(\mathfrak{F},\mathcal{A}\cup\overline{\mathcal{A}}_j^{\varepsilon})$, which is a merger of firms 0 and j when the conditioning statement is satisfied. Now there exists a $\delta > 0$ such that for all $\varepsilon > 0$ the quantity in (39) is at least as large as

$$E_{\mathfrak{F}}[\Delta CS(M_{k'}^{\Pi}) + \delta - E_j^{\mathcal{A}}(\overline{c}_j, \overline{f}_j) | M^*(\mathfrak{F},\mathcal{A}\cup\overline{\mathcal{A}}_j^{\varepsilon}) \in (\mathcal{A}\cup\overline{\mathcal{A}}_j^{\varepsilon})\backslash\mathcal{A}]. \tag{40}$$

As $\varepsilon \to 0$, the quantity in (40) converges to $[\Delta CS(M_{k'}^{\Pi}) - \tau_2] + \delta > 0$, so for small enough $\varepsilon > 0$ (38) implies that this change in the acceptance set is strictly beneficial.

Step 5. For all $j, k \in \mathcal{K}^+$, $j < k$, we must have $\underline{\Delta CS}_j < \underline{\Delta CS}_k$. Suppose otherwise so that for some $j, h \in \mathcal{K}^+$, $h > j$, we have $\underline{\Delta CS}_j \geq \underline{\Delta CS}_h$. Let $k \equiv \arg\min\{h \in \mathcal{K}^+ : h > j \text{ and } \underline{\Delta CS}_j \geq \underline{\Delta CS}_h\}$. Figure 13 shows such a case where $j = 2$ and $k = 3$. Let merger M_k^{CS}'s marginal and fixed costs be $\overline{c}_k^{CS} = \overline{a}_k(\overline{f}_k^{CS})$ and $\overline{f}_k^{CS}$, respectively. Given Step 4, $E_k^{\mathcal{A}}(\overline{a}_k(\overline{f}_k^{CS}), \overline{f}_k^{CS})$ can be written as a weighted average of two conditional expectations:

$$E_{\mathfrak{F}}[\Delta CS(M^*(\mathfrak{F}\backslash M_k,\mathcal{A})) | M_k = (k, \overline{c}_k^{CS}, \overline{f}_k^{CS}),\ M_k = M^*(\mathfrak{F},\mathcal{A}), \text{ and } \Delta\Pi(M^*(\mathfrak{F}\backslash M_k,\mathcal{A})) < \underline{\Delta\Pi}_j] \tag{41}$$

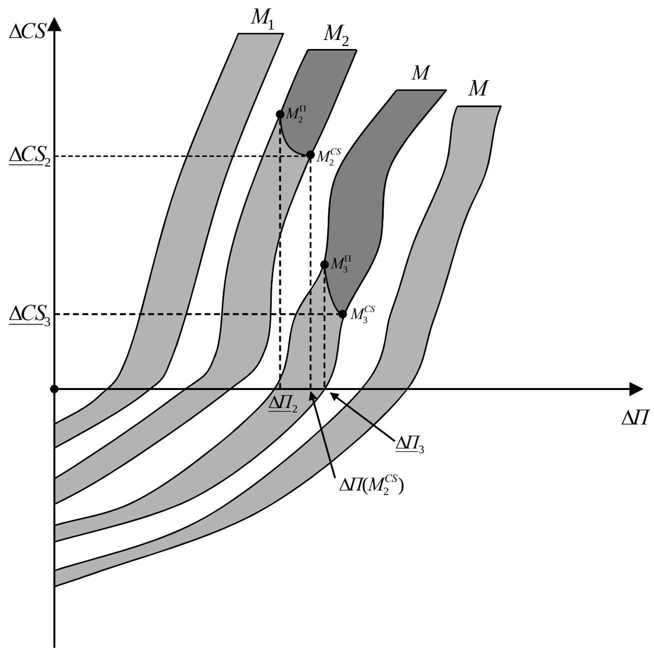

Figure 13: The figure shows the change considered in Step 5 of the proof of Proposition 5

and

$$E_{\mathfrak{F}}\left[\Delta CS(M^*(\mathfrak{F}\backslash M_k,\mathcal{A}))|M_k = (k,\overline{c}_k^{CS},\overline{f}_k^{CS}),\ M_k = M^*(\mathfrak{F},\mathcal{A}),\right.$$
$$\left.\text{and } \Delta\Pi(M^*(\mathfrak{F}\backslash M_k,\mathcal{A})) \in [\underline{\Delta\Pi}_j, \Delta\Pi(M_k)]\right]. \tag{42}$$

Now the term in (41) equals $E_k^{\mathcal{A}}(\overline{a}_k(\overline{f}_k^{\Pi}),\overline{f}_k^{\Pi})$, which by Step 3 equals $\Delta CS(M_j^{\Pi})$, which in turn is at least $\underline{\Delta CS}_j$ by definition. On the other hand, the term in (42) strictly exceeds $\underline{\Delta CS}_j$. Together, this implies that $E_k^{\mathcal{A}}(\overline{a}_k(\overline{f}_k^{CS}),\overline{f}_k^{CS}) > \underline{\Delta CS}_j$. Since, by Step 3, we must have $\underline{\Delta CS}_k = E_k^{\mathcal{A}}(\overline{a}_k(\overline{f}_k^{CS}),\overline{f}_k^{CS})$, this contradicts $\underline{\Delta CS}_j \geq \underline{\Delta CS}_k$.

Step 6. The argument proceeds proceeds along the same lines as that in the proof of Proposition 1. □

7.2 Notes on the Aggregative Game Approach

Assumptions. Suppose an unmerged firm i's profit can be written as

$$\pi(\psi_i, c_i; \Psi),$$

where $\psi_i \geq 0$ is firm i's strategic variable, c_i the firm's constant marginal cost, and $\Psi \equiv \sum_j \psi_j$ an aggregator summarizing the "aggregate outcome." The firm's cumulative best reply, $r(\Psi; c_i) \equiv \arg\max_{\psi_i} \pi(\psi_i, c_i; \psi_i + \sum_{j\neq i}\psi_j)$, is assumed to be single-valued and decreasing in marginal cost c_i. Similarly, a merged firm k's profit is given by $2\pi(\psi_k, \overline{c}_k; \Psi)$, and its cumulative best reply, $\overline{r}(\Psi; \overline{c}_k) \equiv$

$\arg\max_{\psi_k} 2\pi(\psi_k, \overline{c}_k; 2\psi_k + \sum_{j\neq 0,k} \psi_j)$, is single-valued and decreasing in $\overline{c}_k$. Consumer surplus, denoted $V(\Psi)$, is an increasing function of the aggregator and does not depend on the composition of the aggregator.

Suppose that there exists a unique stable equilibrium. Let $\psi_i(M_k)$ denote firm i's equilibrium action under market structure M_k, and $\Psi(M_k) \equiv \sum_j \psi_j(M_k)$. Further, suppose that firm i's equilibrium profit can be written as

$$\begin{aligned} g(\psi_i(M_k); \Psi(M_k)) &\equiv \max_{\psi_i} \pi(\psi_i, c_i; \Psi(M_k)) \text{ if firm } i \text{ is unmerged;} \\ g(2\psi_i(M_k); \Psi(M_k)) &\equiv \max_{\psi_i} 2\pi(\psi_i, \overline{c}_i; \Psi(M_k)) \text{ if firm } i = k \text{ is merged.} \end{aligned}$$

The equilibrium profit function g has the following properties: (i) $g(0; \Psi) = 0$; (ii) for $0 \leq \psi_i \leq \Psi$, $g(\psi_i; \Psi)$ is strictly increasing and strictly convex in ψ_i. We assume that a reduction in post-merger marginal cost $\overline{c}_k$ leads to (a) an increase in $\psi_k(M_k)$ and in the aggregate outcome $\Psi(M_k)$; (b) an increase in $\psi_k(M_k)/\Psi(M_k)$ and a decrease in $\psi_j(M_k)/\Psi(M_k)$, $j \neq 0, k$; and (c) an increase in the merged firm's equilibrium profit $g(2\psi_k(M_k), \Psi(M_k))$ and a reduction in any other firm i's equilibrium profit $g(\psi_i(M_k); \Psi(M_k))$.

Our assumptions hold for several textbook models of competition.

Example 3 (Cournot)**.** In the homogeneous goods Cournot model with constant marginal costs, let ψ_i denote the output of plant i. All unmerged firms can be thought of as single-plant firms, whereas a merged firm can be thought of as running two plants at the same marginal cost (producing the same output at both plants). We impose the same assumptions on demand as in the main text. The profit maximization problem of a single-plant firm i with marginal cost c_i can be written as

$$\max_{\psi_i} \left[P(\psi_i + \sum_{j\neq i} \psi_j) - c_i \right] \psi_i.$$

From the first-order condition of profit maximization, $P(\Psi) - c_i + \psi_i P'(\Psi) = 0$, we can write the equilibrium profit under merger M_k as

$$g(\psi_i(M_k); \Psi(M_k)) = -\left[\psi_i(M_k)\right]^2 P'(\Psi(M_k)).$$

The profit maximization problem of a merged firm k with marginal cost $\overline{c}_k$ (and two plants) can be written as

$$\max_{\psi_k} \left[P(2\psi_k + \sum_{j\neq 0,k} \psi_j) - \overline{c}_k \right] 2\psi_k.$$

From the first-order condition of profit maximization, $P(\Psi) - \overline{c}_k + 2\psi_k P'(\Psi) = 0$, so that we can write the merged firm's equilibrium profit under merger M_k as

$$g(2\psi_k(M_k); \Psi(M_k)) = -\left[2\psi_k(M_k)\right]^2 P'(\Psi(M_k)).$$

It can easily be verified that g has all of the required properties (it takes the value of zero if its first argument is zero and is increasing and convex in its first argument) and that a reduction in post-merger marginal cost $\overline{c}_k$ has the posited effects. (The other assumptions were shown to hold in the main text.)

Example 4 (CES)**.** In the CES demand model with price competition, suppose that an unmerged firm produces a single good, while a merged firm produces two goods at the same marginal cost (thus optimally charging the same price for each). Consider first a single-product firm i. The profit maximization problem of a single-product firm i with marginal cost c_i can be written as

$$\max_{\psi_i}[\psi_i^{-1/\lambda} - c_i]\frac{\psi_i^{(\lambda+1)/\lambda}}{\psi_i + \sum_{j\neq i}\psi_j}.$$

From the first-order condition of profit maximization,

$$-\Psi + \left[\psi_i^{-1/\lambda} - c_i\right]\psi_i^{(\lambda+1)/\lambda}\left\{\frac{(\lambda+1)\Psi}{\psi_i} - \lambda\right\} = 0,$$

it can be seen that there is a unique cumulative best reply $r(\Psi; c_i)$ and that it is decreasing in the firm's marginal cost c_i. We can write the firm's equilibrium profit under merger M_k as

$$g\left(\psi_i(M_k); \Psi(M_k)\right) \equiv \left\{\frac{(\lambda+1)\Psi(M_k)}{\psi_i(M_k)} - \lambda\right\}^{-1}.$$

Consider now the merged firm k and suppose the firm produces two products at marginal cost $\overline{c}_k$. The profit maximization problem can be written as

$$\max_{\psi_k} 2[\psi_k^{-1/\lambda} - \overline{c}_k]\frac{\psi_k^{(\lambda+1)/\lambda}}{2\psi_k + \sum_{j\neq 0,k}\psi_j}.$$

(It can easily be verified that the firm optimally chooses the same value of ψ_k for each one of its two products.) From the first-order condition,

$$-\Psi + \left[\psi_k^{-1/\lambda} - \overline{c}_k\right]\psi_k^{(\lambda+1)/\lambda}\left\{\frac{(\lambda+1)\Psi}{\psi_k} - 2\lambda\right\} = 0,$$

it can be seen that there is a unique cumulative best reply $\overline{r}(\Psi; \overline{c}_k)$ and that it is decreasing in $\overline{c}_k$. We can write the merged firm's equilibrium profit under merger M_k as

$$g\left(2\psi_k(M_k); \Psi(M_k)\right) \equiv \left\{\frac{(\lambda+1)\Psi(M_k)}{2\psi_k(M_k)} - \lambda\right\}^{-1}.$$

It can easily be verified that our assumptions hold in the CES model. In particular, there exists a unique equilibrium and this equilibrium is stable.[21] Moreover, the equilibrium profit function g has all

[21] From the first-order condition for profit maximization, we obtain that $dr(\Psi; c_i)/d\Psi$ can be written as a decreasing and convex function of $\beta_i \equiv \Psi/r(\Psi; c_i)$:

$$\frac{dr(\Psi; c_i)}{d\Psi} = \frac{\lambda}{(\lambda+1)\beta_i(\beta_i - 1) + \lambda}.$$

This derivative attains its maximum of 1 if firm i is the only active firm (i.e., $r(\Psi; c_i) = \Psi$). [Similarly, for a merged firm M_k, we have

$$\frac{d2\overline{r}(\Psi; \overline{c}_k)}{d\Psi} = \frac{\lambda}{(\lambda+1)\overline{\beta}_k(\overline{\beta}_k - 1) + \lambda},$$

where $\overline{\beta}_k \equiv \Psi/\left[2\overline{r}(\Psi; \overline{c}_k)\right]$.] It follows that $\sum_i dr(\Psi; c_i)/d\Psi < 1$ [resp. $\sum_{i\neq 0,k} dr(\Psi; c_i)/d\Psi + 2\overline{r}(\Psi; \overline{c}_k)/d\Psi < 1$ after merger M_k] in any equilibrium with more than one active firm. Hence, any equilibrium must be stable. Moreover, as $r(0; c_i) \geq 0$ [resp. $\overline{r}(0; \overline{c}_k) \geq 0$] and $r(\Psi; c_i) = 0$ [resp. $\overline{r}(\Psi; \overline{c}_k) = 0$] for Ψ sufficiently large, this implies that there exists a unique Ψ that is consistent with equilibrium in the sense that $\Psi - \sum_i r(\Psi; c_i) = 0$ [resp. $\Psi - \sum_{i\neq 0,k} r(\Psi; c_i) - 2\overline{r}(\Psi; \overline{c}_k) = 0$ after merger M_k].

of the required properties (it takes the value of zero if its first argument is zero and is increasing and convex in its first argument). Consider a reduction in post-merger marginal cost $\overline{c}_k$. Since $\overline{r}(\Psi;\overline{c}_k)$ is decreasing in $\overline{c}_k$ and since $r(\Psi;c_i)$ and $\overline{r}(\Psi;\overline{c}_k)$ are increasing in Ψ, and since equilibrium is stable, the reduction in $\overline{c}_k$ induces a higher value of $\Psi = 2\overline{r}(\Psi;\overline{c}_k) + \sum_{i\neq 0,k} r(\Psi;c_i)$. Rewrite the first-order condition of an unmerged firm i:

$$-1 + \left[1 - c_i \left[r(\Psi;c_i)\right]^{1/\lambda}\right] \left\{(\lambda+1) - \lambda \frac{r(\Psi;c_i)}{\Psi}\right\} = 0.$$

As the induced increase in Ψ induces an increase in $r(\Psi;c_i)$ (i.e., prices are strategic complements), the ratio $r(\Psi;c_i)/\Psi$ must fall as otherwise the l.h.s. of the first-order condition would decrease. But as

$$\frac{2\overline{r}(\Psi;\overline{c}_k)}{\Psi} + \frac{\sum_{i\neq 0,k} r(\Psi;c_i)}{\Psi} = 1,$$

it follows that the same ratio for the merged firm, $\overline{r}(\Psi;\overline{c}_k)/\Psi$, must increase. From the expression for the equilibrium profits, we thus obtain that the profit of the merged firm, $g\left(2\overline{r}(\Psi(M_k);\overline{c}_k);\Psi(M_k)\right)$, increases and that of any unmerged firm i, $g\left(r(\Psi(M_k);c_i);\Psi(M_k)\right)$, decreases.

Example 5 (Multinomial Logit)**.** In the multinomial logit demand model with price competition, suppose that an unmerged firm produces a single good, while a merged firm produces two goods at the same marginal cost (and optimally charging the same price for each). Consider first a single-product firm i. The profit maximization problem of a single-product firm i with marginal cost c_i can be written as

$$\max_{\psi_i} \left[a - \mu \ln \psi_i - c_i\right] \frac{\psi_i}{\psi_i + \sum_{j\neq i} \psi_j}$$

From the first-order condition of profit maximization,

$$\left\{-\mu + a - \mu \ln \psi_i - c_i\right\} \Psi - \left[a - \mu \ln \psi_i - c_i\right] \psi_i = 0,$$

it can be seen that there is a unique cumulative best reply $r(\Psi;c_i)$ and that it is decreasing in the firm's marginal cost c_i. Firm i's equilibrium profit under merger M_k can be written as

$$g(\psi_i(M_k);\Psi(M_k)) = \mu \left\{\frac{\Psi(M_k)}{\psi_i(M_k)} - 1\right\}^{-1}.$$

Consider now the merged firm k and suppose the firm produces two products at marginal cost $\overline{c}_k$. The profit maximization problem can be written as

$$\max_{\psi_k} 2\left[a - \mu \ln \psi_k - \overline{c}_k\right] \frac{\psi_k}{2\psi_k + \sum_{j\neq 0,k} \psi_j}.$$

(It can easily be verified that the firm optimally chooses the same value of ψ_k for each one of its two products.) From the merged firm's first-order condition of profit maximization,

$$\left\{-\mu + a - \mu \ln \psi_k - \overline{c}_k\right\} \Psi - 2\left[a - \mu \ln \psi_k - \overline{c}_k\right] \psi_k = 0,$$

it can be seen that there is a unique cumulative best reply $\overline{r}(\Psi;\overline{c}_k)$ and that it is decreasing in $\overline{c}_k$. Firm k's equilibrium profit under merger M_k can be written as

$$g(2\psi_k(M_k);\Psi(M_k)) = \mu \left\{\frac{\Psi(M_k)}{2\psi_k} - 1\right\}^{-1}.$$

It can easily be verified that our assumptions hold in the multinomial logit model. In particular, there exists a unique equilibrium and this equilibrium is stable.[22] Moreover, the equilibrium profit function g has all of the required properties (it takes the value of zero if its first argument is zero and is increasing and convex in its first argument). Consider a reduction in post-merger marginal cost $\overline{c}_k$. Since $\overline{r}(\Psi;\overline{c}_k)$ is decreasing in $\overline{c}_k$ and since $r(\Psi;c_i)$ and $\overline{r}(\Psi;\overline{c}_k)$ are increasing in Ψ, and since equilibrium is stable, the reduction in $\overline{c}_k$ induces a higher value of $\Psi = 2\overline{r}(\Psi;\overline{c}_k) + \sum_{i\neq 0,k} r(\Psi;c_i)$. Rewrite the first-order condition of an unmerged firm i:

$$-1 + \left[1 - c_i \left[r(\Psi;c_i)\right]^{1/\lambda}\right] \left\{(\lambda+1) - \lambda \frac{r(\Psi;c_i)}{\Psi}\right\} = 0.$$

As the induced increase in Ψ induces an increase in $r(\Psi;c_i)$ (i.e., prices are strategic complements), the ratio $r(\Psi;c_i)/\Psi$ must fall as otherwise the l.h.s. of the first-order condition would decrease. But as

$$\frac{-\mu + a - \mu \ln r(\Psi;c_i) - c_i}{a - \mu \ln r(\Psi;c_i) - c_i} = \frac{r(\Psi;c_i)}{\Psi}.$$

it follows that the same ratio for the merged firm, $\overline{r}(\Psi;\overline{c}_k)/\Psi$, must increase. From the expression for the equilibrium profits, we thus obtain that the profit of the merged firm, $g\left(2\overline{r}(\Psi(M_k);\overline{c}_k);\Psi(M_k)\right)$, increases and that of any unmerged firm i, $g\left(r(\Psi(M_k);c_i);\Psi(M_k)\right)$, decreases.

Results. Let $\psi_i^0 \equiv \psi_i(M_0)$ and $\Psi^0 \equiv \Psi(M_0)$, and note that, as consumer surplus $V(\Psi)$ is strictly increasing in Ψ, merger M_k is CS-neutral if $\Psi(M_k) = \Psi^0$; it is CS-increasing if $\Psi(M_k) > \Psi^0$, and CS-decreasing if $\Psi(M_k) < \Psi^0$.

Lemma 9. *Merger M_k is CS-neutral if $2\psi_k(M_k) = \psi_0^0 + \psi_k^0$, CS-increasing if $2\psi_k(M_k) > \psi_0^0 + \psi_k^0$, and CS-decreasing if $2\psi_k(M_k) < \psi_0^0 + \psi_k^0$.*

Proof. Suppose merger M_k is CS-neutral. Then, $\Psi(M_k) = \Psi^0$. From the profit maximization problem of any firm i not involved in the merger, it follows that $\psi_i(M_k) = r\left(\Psi(M_k);c_i\right) = r\left(\Psi^0;c_i\right) = \psi_i^0$. Hence, we must have $2\psi_k(M_k) = \psi_0^0 + \psi_k^0$. The claim then follows from the observation that consumer surplus is increasing in Ψ and that the equilibrium is stable. □

Lemma 10. *If merger M_k is CS-neutral, it raises the joint profit of the merging firms as well as aggregate profit.*

Proof. It is immediate to see that the profit of any firm not involved in the merger remains unchanged as Ψ remains unchanged. It thus remains to show that

$$g(2\psi_k(M_k);\Psi(M_k)) > g\left(\psi_0^0;\Psi^0\right) + g\left(\psi_k^0;\Psi^0\right).$$

[22] From the first-order condition for profit maximization, we obtain that $dr(\Psi;c_i)/d\Psi$ can be written as a decreasing and convex function of $\beta_i \equiv \Psi/r(\Psi;c_i)$:

$$\frac{dr(\Psi;c_i)}{d\Psi} = \frac{1}{\beta_i(\beta_i - 1) + 1}.$$

This derivative attains its maximum of 1 if firm i is the only active firm (i.e., $r(\Psi;c_i) = \Psi$). [Similarly, for a merged firm M_k, we have

$$\frac{d2\overline{r}(\Psi;\overline{c}_k)}{d\Psi} = \frac{1}{\overline{\beta}_k(\overline{\beta}_k - 1) + 1},$$

where $\overline{\beta}_k \equiv \Psi/\left[2\overline{r}(\Psi;\overline{c}_k)\right]$.] It follows that $\sum_i dr(\Psi;c_i)/d\Psi < 1$ [resp. $\sum_{i\neq 0,k} dr(\Psi;c_i)/d\Psi + 2\overline{r}(\Psi;\overline{c}_k)/d\Psi < 1$ after merger M_k] in any equilibrium with more than one active firm. Hence, any equilibrium must be stable. Moreover, as $r(0;c_i) \geq 0$ [resp. $\overline{r}(0;\overline{c}_k) \geq 0$] and $r(\Psi;c_i) = 0$ [resp. $\overline{r}(\Psi;\overline{c}_k) = 0$] for Ψ sufficiently large, this implies that there exists a unique Ψ that is consistent with equilibrium in the sense that $\Psi - \sum_i r(\Psi;c_i) = 0$ [resp. $\Psi - \sum_{i\neq 0,k} r(\Psi;c_i) - 2\overline{r}(\Psi;\overline{c}_k) = 0$ after merger M_k].

But as M_k is CS-neutral, we have $\Psi(M_k) = \Psi^0$ and $2\psi_k(M_k) = \psi_0^0 + \psi_k^0$. The above inequality can thus be rewritten as

$$g\left(\psi_0^0 + \psi_k^0; \Psi^0\right) > g\left(\psi_0^0; \Psi^0\right) + g\left(\psi_k^0; \Psi^0\right).$$

But this follows from the assumed properties of the function g. □

As a reduction in post-merger marginal cost increases the merged firm's profit, any CS-nondecreasing merger is profitable. As in the Cournot model with efficient bargaining (Section 5.1.1), we impose the following assumption:

Assumption 4. *If merger M_k, $k \geq 1$, is CS-nondecreasing, then reducing its post-merger marginal cost $\overline{c}_k$ increases the aggregate profit* $\left[g(2\psi_k(M_k); \Psi(M_k)) + \sum_{i \in \mathcal{N} \setminus \{0,k\}} g(\psi_i(M_k); \Psi(M_k))\right]$.

In the CES and multinomial logit models (and, as we have seen before, in the Cournot model), a sufficient condition for this assumption to hold is that pre-merger cost differences are not too large so that for every merger M_k, $\left(\psi_0^0 + \psi_k^0\right)/\Psi^0 > \max_{i \neq 0,k} \psi_i^0/\Psi^0$, i.e., the sum of the pre-merger shares of the merger partners exceeds the pre-merger share of the largest nonmerging firm.

Example 6 (CES)**.** In the CES model, if pre-merger marginal cost differences are not too large so that $\left(\psi_0^0 + \psi_k^0\right)/\Psi^0 > \max_{i \neq 0,k} \psi_i^0/\Psi^0$, then the reduction in post-merger marginal cost $\overline{c}_k$ following a CS-nondecreasing merger M_k increases aggregate profit. To see this, note that from the argument given in our exposition of the CES model above, the reduction in $\overline{c}_k$ induces a change from ψ_i/Ψ to $(\psi_i/\Psi - \Delta_i)$, $i \neq 0, k$, $\Delta_i > 0$, and from $2\psi_k/\Psi$ to $(2\psi_k/\Psi + \sum_{i \neq 0,k} \Delta_i)$. It thus suffices to show that the joint profit of the merged firm k and any other firm i,

$$h_i(\Delta) \equiv \left\{\frac{\sigma_k + \Delta}{(\lambda+1) - \lambda(\sigma_k + \Delta)}\right\} + \left\{\frac{\sigma_i - \Delta}{(\lambda+1) - \lambda(\sigma_i - \Delta)}\right\},$$

where $\Delta \in [0, \Delta_i]$, $\sigma_i = \psi_i/\Psi$ and $2\psi_k/\Psi \leq \sigma_k \leq 2\psi_k/\Psi + \sum_{j \neq 0,k} \Delta_j$, is increasing in Δ. But this holds as we have

$$h_i'(\Delta) \equiv \frac{\lambda+1}{\left[(\lambda+1) - \lambda(\sigma_k + \Delta)\right]^2} - \frac{\lambda+1}{\left[(\lambda+1) - \lambda(\sigma_i - \Delta)\right]^2} > 0,$$

where the inequality follows since $\psi_0^0 + \psi_k^0 > \psi_i^0$ implies that $\sigma_k > \sigma_i$ for any CS-nondecreasing merger M_k.

Example 7 (Multinomial Logit)**.** In the multinomial logit model, if pre-merger marginal cost differences are not too large so that $\left(\psi_0^0 + \psi_k^0\right)/\Psi^0 > \max_{i \neq 0,k} \psi_i^0/\Psi^0$, then the reduction in post-merger marginal cost $\overline{c}_k$ following a CS-nondecreasing merger M_k increases aggregate profit. To see this, note that from the argument given in our exposition of the multinomial logit model above, the reduction in $\overline{c}_k$ induces a change from ψ_i/Ψ to $(\psi_i/\Psi - \Delta_i)$, $i \neq 0, k$, $\Delta_i > 0$, and from $2\psi_k/\Psi$ to $(2\psi_k/\Psi + \sum_{i \neq 0,k} \Delta_i)$. It thus suffices to show that the joint profit of the merged firm k and any other firm i,

$$h_i(\Delta) \equiv \mu\left\{\frac{\sigma_k + \Delta}{1 - (\sigma_k + \Delta)}\right\} + \mu\left\{\frac{\sigma_i - \Delta}{1 - (\sigma_i - \Delta)}\right\},$$

where $\Delta \in [0, \Delta_i]$, $\sigma_i = \psi_i/\Psi$ and $2\psi_k/\Psi \leq \sigma_k \leq 2\psi_k/\Psi + \sum_{j \neq 0,k} \Delta_j$, is increasing in Δ. But this holds as we have

$$h_i'(\Delta) \equiv \frac{\mu}{\left[1 - (\sigma_k + \Delta)\right]^2} - \frac{\mu}{\left[1 - (\sigma_i - \Delta)\right]^2} > 0,$$

where the inequality follows since $\psi_0^0 + \psi_k^0 > \psi_i^0$ implies that $\sigma_k > \sigma_i$ for any CS-nondecreasing merger M_k.

We are now in the position to extend Lemma 4 to this larger class of models:

Lemma 11. *Suppose mergers M_j and M_k, $k > j$, induce the same nonnegative change in consumer surplus so that $\Psi(M_j) = \Psi(M_k) \geq \Psi^0$. Then, the larger merger M_k induces a greater increase in aggregate profit than the smaller merger M_j.*

Proof. As the aggregate outcome Ψ is the same under both mergers, the profit of each firm not participating in either merger is also the same under both mergers. We thus only need to show that

$$g(2\psi_k(M_k);\overline{\Psi}) + g(\psi_j(M_k);\overline{\Psi}) > g(2\psi_j(M_j);\overline{\Psi}) + g(\psi_k(M_j);\overline{\Psi}),$$

where $\overline{\Psi} \equiv \Psi(M_j) = \Psi(M_k)$ is the common aggregate outcome after each of the two alternative mergers. As $\Psi(M_j) = \Psi(M_k)$, we must have

$$2\psi_k(M_k) + \psi_j(M_k) = 2\psi_j(M_j) + \psi_k(M_j).$$

Now, as $c_j > c_k$ and as $\Psi(M_j) = \Psi(M_k)$, we obtain (from the assumption that a firm's cumulative best reply is decreasing in its marginal cost) that

$$\psi_j(M_k) < \psi_k(M_j),$$

implying that

$$2\psi_k(M_k) > 2\psi_j(M_j).$$

Next, note that as a CS-nondecreasing merger increases the profit of the merging firms and reduces everyone else's profit, we have

$$\begin{aligned} g(2\psi_k(M_k), \Psi(M_k)) &> g(\psi_k^0, \Psi^0) \\ &\geq g(\psi_k(M_j), \Psi(M_j)). \end{aligned}$$

As $\Psi(M_k) = \Psi(M_j)$ and as g is strictly increasing in its first argument, this implies that

$$2\psi_k(M_k) > \psi_k(M_j).$$

Using the same type of argument, we also have

$$2\psi_j(M_j) > \psi_j(M_k).$$

We have thus shown that

$$2\psi_k(M_k) > \max\left\{2\psi_j(M_j), \psi_k(M_j)\right\} \geq \min\left\{2\psi_j(M_j), \psi_k(M_j)\right\} > \psi_j(M_k).$$

But since $2\psi_k(M_k) + \psi_j(M_k) = 2\psi_j(M_j) + \psi_k(M_j)$ and since g is strictly convex in its first argument, this implies that

$$g(2\psi_k(M_k);\overline{\Psi}) + g(\psi_j(M_k);\overline{\Psi}) > g(2\psi_j(M_j);\overline{\Psi}) + g(\psi_k(M_j);\overline{\Psi}).$$

□

Finally, note that if $\left|\overline{\Psi} - \Psi^0\right|$ is sufficiently small, where $\overline{\Psi} \equiv \Psi(M_j) = \Psi(M_k) \geq \Psi^0$, then the lemma also implies that the larger merger M_k induces a larger increase in the bilateral profit change than the smaller merger M_j. (This follows from the fact that if both mergers are CS-neutral, then the induced bilateral profit change is equal to the induced aggregate profit change.)

References

[1] Alonso, R. and N. Matouschek (2008), "Optimal Delegation," *Review of Economic Studies* 75(1): 259-293.

[2] Armstrong, M. and J. Vickers (2010), "A Model of Delegated Project Choice," *Econometrica* 78(1) 213-44.

[3] Bernheim, B. D., and M. D. Whinston (1986), "Menu Auctions, Resource Allocation, and Economic Influence," *Quarterly Journal of Economics* 101(1): 1-31.

[4] Besanko, D. and D. Spulber (1993), "Contested Mergers and Equilibrium Antitrust Policy," *Journal of Law, Economics, and Organization* 9(1): 1-29.

[5] Che, Y.-K., W. Dessein, and N. Kartik (2010), "Pandering to Persuade," CEPR Discussion Paper DP7970.

[6] Dubey, P., O. Haimanko, and A. Zapechelnyuk (2006), "Strategic Complements and Substitutes, and Potential Games," *Games and Economic Behavior* 54(1): 77-94.

[7] Farrell, J. and C. Shapiro (1990), "Horizontal Mergers: An Equilibrium Analysis," *American Economic Review* 80(1): 107-26.

[8] Holmström, B. (1984), "On the Theory of Delegation," in *Bayesian Models in Economic Theory*, ed. by M. Boyer, and R. Kihlstrom. North-Holland, New York.

[9] Jehiel, P., B. Moldovanu, and E. Stacchetti (1996), "How (Not) to Sell Nuclear Weapons," *American Economic Review* 86(4): 814-29.

[10] Lyons, B. R. (2003), "Could Politicians be More Right Than Economists? A Theory of Merger Standards," European University Institute working paper #2003/14.

[11] McAfee, R. P. and M. A. Williams (1992), "Horizontal Mergers and Antitrust Policy," *Journal of Industrial Economics* 40(2): 181-7.

[12] Nocke, V. and M. D. Whinston (2010), "Dynamic Merger Review," *Journal of Political Economy* 118(6), 1200-1251.

[13] Ottaviani, M. and A. Wickelgren (2009), "Policy Timing Under Uncertainty: Ex Ante versus Ex Post Merger Control," working paper.

[14] Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model," *Econometrica* 50(1): 97-109

[15] Segal, I. (1999), "Contracting with Externalities," *Quarterly Journal of Economics* 114(2): 337-88.

[16] Williamson, O.E. (1968), "Economies as an Antitrust Defense: The Welfare Tradeoffs," *American Economic Review* 58(1): 18-36.