

## SCREENING IN VERTICAL OLIGOPOLIES

HECTOR CHADE

Department of Economics, Arizona State University

JEROEN SWINKELS

Kellogg School of Management, Northwestern University

A finite number of vertically differentiated firms simultaneously compete for and screen agents with private information about their payoffs. In equilibrium, higher firms serve higher types. Each firm distorts the allocation downward from the efficient level on types below a threshold, but upward above. While payoffs in this game are neither quasi-concave nor continuous, if firms are sufficiently differentiated, then any strategy profile that satisfies a simple set of necessary conditions is a pure-strategy equilibrium, and an equilibrium exists. A mixed-strategy equilibrium exists even when firms are less differentiated. The welfare effects of private information are drastically different than under monopoly. The equilibrium approaches the competitive limit quickly as entry costs grow small. We solve the problem of a multi-plant firm facing a type-dependent outside option and use this to study the effect of mergers.

KEYWORDS: Adverse selection, screening, quality distortions, oligopoly, incentive compatibility, positive sorting, vertical differentiation, merger analysis, competitive limit, equilibrium existence.

### 1. INTRODUCTION

SCREENING IS CENTRAL TO LABOR and product markets. In [Mussa and Rosen \(1978\)](#) and [Maskin and Riley \(1984\)](#), a monopolist screens a consumer with private information about his valuation. In [Rothschild and Stiglitz \(1976\)](#) and variations thereof, identical insurance companies competitively screen consumers. Most markets do not fall at these extremes of monopoly or perfect competition. Instead, a small number of heterogeneous firms both compete for and screen their customers. The quality and price of any given Saint Laurent handbag affects the sales of its other handbags, but also affects how it competes with the artisans of Hermès above them and deep supply chains of Coach below. On any given route, Delta Airlines offers its customers a multitude of quality levels, but may compete with Southwest below and a private jet firm above. Consumer-packaged-goods firms sell at multiple quality and price points, but in an oligopolistic environment. Consulting firms screen their workers into appropriate roles, but compete for talent. The incumbent cable provider can provide fast internet access more easily than can the legacy telephone company, but each screens their own customers. A firm offering a dedicated connection to the internet backbone can offer even faster speeds and, for financial firms, a firm offering a server collocated with those of the financial exchange can offer faster speeds still.

---

Hector Chade: [hector.chade@asu.edu](mailto:hector.chade@asu.edu)

Jeroen Swinkels: [j-swinkels@northwestern.edu](mailto:j-swinkels@northwestern.edu)

We thank the four referees for their helpful comments and suggestions. We are also grateful to Meghan Busse, Laura Doval, Bruno Jullien, Andreas Kleiner, Natalia Kovrijnykh, Michael Powell, Luis Rayo, and, especially, John Conlon, along with seminar participants at Northwestern, 2018 SE, Mexico City, Universidad Alberto Hurtado, Chile, CIDE, Mexico, University of Georgia, Universitat Pompeu Fabra, Federal Reserve Bank of Minneapolis, University of Hawaii, Universidad Nacional de Cuyo, Argentina, 2020 Stony Brook Game Theory Conference, and Yale University for their helpful comments.

© 2021 The Authors. *Econometrica* published by John Wiley & Sons Ltd on behalf of The Econometric Society. Jeroen Swinkels is the corresponding author on this paper. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](#) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

The lack of a standard workhorse for oligopolistic screening has hindered progress theoretically and empirically, and leaves important economic questions open. What do equilibria look like? Do our standard intuitions about screening still hold? Who does asymmetric information help or hurt? Is price discrimination pro- or anti-competitive? Does increasing competition lead toward efficiency despite asymmetric information? Are the effects of mergers unambiguous? Do equilibria in pure strategies even exist or are such markets inherently unstable?

This paper takes an important step in filling this gap. Working in a model of vertical differentiation and building on the duopoly model of Biglaiser and Mezzetti (1993), we consider an oligopoly with a finite set of firms facing a continuum of consumers with quasi-linear preferences and willingness to pay for quality in a product market that is increasing in their privately known type. The firms themselves are also ranked, in the sense that a higher indexed firm has a cost of quality that is shallower than that of a lower indexed firm, and crosses it from above, so that the lower firm has an advantage at producing low quality goods and the higher indexed firm has an advantage at producing high quality goods. They are in this sense *vertically* ranked. The model can equally be interpreted as one of consumers with different privately known willingness to pay and firms with differing marginal costs of providing quantity, or as a labor market with workers of differing privately known ability and firms with successively higher marginal values for ability.

Our model is thus directly applicable in settings of vertical differentiation such as internet access speed, or competition on many airline routes. We also expect it to be a significant building block where vertical and horizontal differentiation coexist, as when Saint Laurent also competes horizontally with Bottega Veneta or Delta competes with United.

We provide necessary conditions for equilibrium and show that they are sufficient if firms are sufficiently differentiated. This allows us to prove pure-strategy equilibrium existence and allows easy numerical analysis of how equilibria vary with the underlying structure. We study the welfare effects of asymmetric information, the competitive limit as entry costs grow small, and mergers.

We model this as a simultaneous game among firms who post menus of incentive-compatible contracts. A menu consists of transfer–action pairs or, equivalently, an action and a surplus as a function of the agent’s type. We rule out contracts that condition on the offers of other firms.<sup>1</sup> Having seen the available offers, agents then choose the firm and contract that suits them best, resolving ties across firms equiprobably. We focus primarily on pure-strategy Nash equilibria.

We first derive a set of properties that any pure-strategy equilibrium exhibits.<sup>2</sup> Our model has private values: the type of an agent enters the firm’s profit only through the contract chosen. Using this, we show that firms make *positive profit* on each type served. Any equilibrium also satisfies *no-poaching*: imitating the contract offered by the incumbent to any given type yields negative profit to the imitating firm. Thus, the agent is matched to the firm that creates the most surplus for the action level chosen, but this action will generically be inefficient even given the chosen firm, and a more efficient choice of firm match may exist.

Our model embeds a nontrivial matching problem. We show that any equilibrium entails *positive sorting*: higher firms serve higher types. If firms are not very differentiated,

<sup>1</sup>This is not without loss of generality (Epstein and Peters (1999), Martimort and Stole (2002)), but is economically reasonable in most settings.

<sup>2</sup>As discussed fully below, several of these necessary conditions are closely related to those that Jullien (2000) derives in the case of a single principal who faces a type-dependent participation constraint.

then adjacent firms may also tie on an interval of types at their shared boundary, in which case they will offer a zero-profit contract to those types. Positive sorting under incomplete information highlights the dual role that menus play: screening the types served, but also attracting the right pool of types.

Since firms serve intervals of types, each firm can solve for the optimal interval served and the optimal menu given that interval. Over the relevant interval, the firm's menu satisfies *internal optimality*: actions are pinned down by a condition that generalizes the standard monopoly trade-off between efficiency and information rents. The difference from the standard condition reflects that the firm serves only a segment of the market, and in general faces a binding participation constraint at both the bottom and top of the interval served. Each firm must also satisfy *optimal boundary* conditions reflecting that changing the action of a boundary type alters this type's profit, but also attracts or loses some nearby types.

These conditions yield a clear pattern of distortions. The highest firm distorts all actions downward: lowering the action of a type lowers the information rents of higher types. In turn, the lowest firm distorts effort upward for all types, as the option of being served by someone else binds only for the highest type served, and raising actions lowers the information rents of lower types. For a middle firm, participation binds for both the lowest and highest agent served. The firm can lower the information rents of middle types by distorting the action downward for types below a threshold, and upward above. When firms are sufficiently differentiated, there are action gaps at the boundaries between adjacent firms. Thus, in a labor market setting, higher firms may ask more of their least able worker than the firm below them asks of their most able worker. Similarly, products of certain intermediate qualities are simply not offered.

In 1849, Dupuit argued that a rail company provides roofless carriages in third class to "frighten the rich" (see [Tirole \(1988, p. 150\)](#)), for the full quotation). Consistent with the extant theory, this reduces profits on the poor only a little and helps to sell second-class seats. But Dupuit also argues that first-class passengers receive "superfluous" quality, something more puzzling in the standard theory.<sup>3</sup> However, as our results show, if the rail company competes against higher quality alternatives, then for the rich, an inefficiently high quality can be largely reflected in the price, but the high price again helps to sell second-class seats to the middle class.

We then turn to sufficiency and existence. If firms are differentiated enough—a condition we call *stacking*—then any strategy profile that satisfies positive sorting, internal optimality, and optimal boundaries is (essentially) an equilibrium, and we need not check no-poaching. One implication of this is that an equilibrium is characterized by a numerically tractable set of equations, which we exploit for examples and exploration. The second crucial implication of this simplification is that an equilibrium in pure strategies *exists*.<sup>4</sup>

Sufficiency and pure strategy existence are central. They are fundamental for applications, and the proof is novel and of broader scope. One challenge is that payoffs are discontinuous: a firm offering less surplus than its competitors never wins, while one that offers more does so always. A deeper problem is that two strategies for a given player may earn the same payoff, but serve different sets of agents, and so their convex combination, which will serve yet a different set of agents, will relate to neither of them tractably. This lack of quasi-concavity makes sufficiency both surprising and nontrivial, and complicates the use of off-the-shelf existence results. The key is to reparameterize our problem

<sup>3</sup>Indeed, see the paragraph in [Tirole \(1988\)](#) immediately after the quote.

<sup>4</sup>Existence of a sensible pure-strategy equilibrium when firms are less differentiated is an open question.

into the much lower dimensional problem of choosing optimal boundaries given that one acts optimally on the interval of types served. While the “topography” of payoffs remains complicated, we establish the existence of a unique optimum characterized by the optimal boundary conditions, and use this to establish sufficiency and existence.

We next compare our model to one with complete information. In a monopoly, complete information hurts the agent (and helps the firm) by destroying information rents. Here, we have a surprising partial reversal. Information rents again disappear, but firms can now compete more aggressively for types served by another firm without attracting their own types, and so the outside option improves: each type now receives the surplus that the second most capable firm for his type can provide. This generates intervals of types who prefer complete information. Indeed, we show that *all* types may prefer complete information.

This result points to an interesting trade-off in a world where firms have increasingly good data on their customers. When there is a monopolist provider (as might be argued for Amazon in many segments), then regulations banning them from charging different people different prices are pro-consumer, effectively restoring asymmetric information. But if there is a capable competitor in an adjacent segment (perhaps Walmart in some e-commerce segments), then allowing firms to tailor offers may incentivize them to compete aggressively on a broader array of customers.

Three forces pull the equilibrium surplus of any given type down from the competitive equilibrium: the action will be distorted from efficiency for the firm to which the type is matched, the type and the firm may be mismatched, and the firm to whom the type is matched earns rents. A natural question is what happens with these forces as the market grows large. To this end, we study a version of our model where firms can enter the market at a fixed cost and choose their technology. As the fixed cost shrinks, the number of firms,  $N$ , grows, and we approach a competitive limit. The profit per type and loss in consumer surplus are on the order of  $1/N^2$ , where the first  $N$  captures the extent of differentiation, while the second captures that when market share is small, the trade-off in chasing extra market share is favorable.

We then consider the effects of mergers in our setting. We first analyze the benchmark case where all the firms merge to form a monopolist with multiple technologies, and we show that both internal optimality and the optimal boundaries condition hold for the merged firm. In addition, there are internal boundary conditions among the constituent technologies which internalize the externality that a type gained by one technology is lost by an adjacent one. Compared to the oligopoly case, and even if forced to serve the same set of types, the multi-firm monopoly will reconfigure its action profile to reduce the information rents of its interior types. Moreover, it will wish to shed some types at both ends. The proof of this result is nontrivial, because the firm is simultaneously adjusting its action profile. The policy takeaway is that to protect customers post-merger, it is not enough to insist that the firm does not shed customers: the wellbeing of interior customers must also be addressed.

Clearly, the same results hold if only a subset of adjacent firms merge and we hold fixed the behavior of other firms. But in equilibrium, nonmerged firms also adjust their behavior post-merger. It is intuitive, and true in all of our numerical examples, that post-merger, all firms will offer a worse deal, since the initial impact of the merger is for the merged firm to shed market share. There is a trade-off: the merged firm will also typically move its actions at its boundary types closer to those of the competitors so as to decrease information rents of interior types. This makes it tempting for adjacent firms to *raise* the surplus they offer and pick up extra market share, and prevents a clean theoretical prediction, except in special cases, one of which we explore.

Finally, we return to the issue of existence when stacking fails. We show that under efficient tie-breaking, an equilibrium in mixed strategies exists. The proof of this result contains an idea that may be useful in other applications.

## 2. RELATED LITERATURE

This paper relates to the immense literature on principal–agent models with screening/adverse selection (see [Mussa and Rosen \(1978\)](#) and [Maskin and Riley \(1984\)](#), and see [Laffont and Martimort \(2002\)](#) for a survey). It is more related to the small literature on oligopoly and price discrimination under adverse selection (see [Stole \(2007\)](#) for a survey). [Champsaur and Rochet \(1989\)](#) analyze a two-stage game where two firms first choose intervals of qualities they can produce and then offer price schedules to consumers. Since firms can cede parts of the market before price competition takes place, the economics are very different. [Spulber \(1989\)](#), working in a [Salop \(1979\)](#) model of horizontal differentiation, considers screening on quantities.<sup>5</sup> The surplus schedule is as in monopoly, with intercept determined by competition. [Stole \(1995\)](#) analyzes oligopoly with screening. In the relevant case, the vertical dimension is private information while the horizontal one is known. Critically, providing quality costs the same to each firm, and so each firm serves all close-by customers regardless of their vertical type. The matching patterns of heterogeneous firms and agents is at the heart of our analysis.

[Biglaiser and Mezzetti \(1993\)](#) analyze a closely related version of our setup with two firms. They find, as we do, that matching is assortative, that the lowest firm distorts uniformly upward, and that there may be a region of ties. But much of economic interest requires that we move beyond two firms, especially regarding the pattern of equilibrium distortions, welfare effects, and applications such as mergers and the competitive limit. Our sufficiency and existence analysis also shows that finding an equilibrium reduces to solving a tractable (analytically and numerically) system of local optimality conditions. Another difference is that in their paper ties are broken in favor of a firm that gains the most from that type, which tames payoff discontinuities. In most of our analysis, we assume the equiprobable rule (or, more generally, each tying firm wins with positive probability) and need to tackle payoff discontinuities head on.

[Jullien \(2000\)](#) provides a sophisticated analysis of a principal–agent model with type-dependent reservation utility, where both upward and downward distortions can emerge (generalizing [Maggi and Rodriguez-Clare \(1995\)](#)). Our firms face an outside option driven by competitors, so some of our conditions have a close relative in Jullien. A significant technical difference is that a firm that matches the outside option always wins in Jullien's model, but this cannot hold for all firms simultaneously in oligopoly. Those of our necessary conditions which derive from equilibrium, including no-poaching and positive sorting, are novel. And since our model endogenizes the agents' reservation utility, we provide a more clear-cut prediction of the equilibrium distortions.

In Theorem 4, Jullien shows conditions under which his necessary conditions are also sufficient with full participation, while Section 4 extends the model to cases without full participation. But one cannot apply the sufficiency part of Theorem 4 and the ideas of Section 4 at the same time, since sufficiency requires that the benefit to the firm from the action is concave, while Section 4 adds an artificial technology that mimics the outside

---

<sup>5</sup>For an interesting recent contribution of screening with horizontal differentiation in a Hotelling model of a labor market with two firms, binary types, and multitasking, see [Benabou and Tirole \(2016\)](#).

option of the agent. As a maximum of two functions, the technology constructed in this way fails concavity.<sup>6</sup>

Our paper also relates to many-to-one matching problems with transfers, as in Crawford and Knoer (1981) and Kelso and Crawford (1982). A recent paper on matching models with “large” firms (and complete information) is Eeckhout and Kircher (2018).<sup>7</sup> Finally, there is a large literature on competitive markets with adverse selection in the tradition of Rothschild and Stiglitz (1976), including recent contributions featuring search frictions, as in Guerrieri, Shimer, and Wright (2010) and Lester, Shourideh, Venkateswaran, and Zetlin-Jones (2018).

### 3. THE MODEL

There is a unit measure of agents (workers or customers) and there are  $N$  principals (firms). Agents have type  $\theta \in [0, 1]$  with cumulative distribution function (cdf)  $H$  with strictly positive and  $\mathcal{C}^1$  density  $h$ .<sup>8</sup> We assume  $H$  and  $1 - H$  are strictly log-concave.<sup>9</sup>

The agent chooses an action  $a \geq 0$ . The value before transfers of action  $a$  to firm  $n$  is  $\mathcal{V}^n(a)$ , and the value to the agent of type  $\theta$  of action  $a$  is  $\mathcal{V}(a, \theta)$ . These objects are  $\mathcal{C}^2$ , and  $\mathcal{V}^n(a)$  is strictly supermodular in  $n$  and  $a$ . For simplicity, we assume that  $\mathcal{V}$  can be written as  $\mathcal{V}(a, \theta) = \mathcal{U}(a) + a\theta$ . Having  $\mathcal{V}_\theta = a$  adds substantial tractability and we believe does not subtract significantly from the economics of the situation. Payoffs given action  $a$  and transfer  $t$  are  $\mathcal{V}^n(a) + t$  and  $\mathcal{V}(a, \theta) - t$ , where  $t$  would typically be positive in a product market, where the agent is a customer, and typically be negative in a labor market, where the agent is a worker. Note that  $\mathcal{V}^n(a)$  does not depend on the identity of the agent,  $\theta$ , and, similarly,  $\mathcal{V}(a, \theta)$  does not depend on the identity of the firm, capturing that our model is of private values and has no horizontal element. For simplicity, for much of the paper we assume that the agent has no outside option beyond the offers of the various firms.<sup>10</sup> To zero in on competition under adverse selection, we assume that the action is observable, thus ruling out moral hazard. Firms do not have capacity constraints and their technology is additively separable across agents.

Define  $V^n(a) = \mathcal{V}^n(a) + \mathcal{U}(a)$ , so that  $V^n(a) + a\theta$  is the match surplus between firm  $n$  and type  $\theta$  with action  $a$ . For each  $n$ ,  $V^n$  is strictly concave, and we also assume that  $V_a^n(0) > 1/h(0)$  and  $\lim_{a \rightarrow \infty} V_a^n(a) < -1 - (1/h(1))$ , which will ensure that implemented actions are interior.

**EXAMPLE 1—A product market with quality differentiation:** Let  $\mathcal{V}(a, \theta) = \sqrt{\rho + a} + a\theta$ , where  $\rho > 0$  is sufficiently small, be the value to the customer of type  $\theta$  of product quality  $a$ . Let  $\mathcal{V}^n(a) = -c^n(a)$ , where  $c^n$  is the cost to firm  $n$  of quality  $a$ , and where  $c^n$  is strictly increasing and convex in  $a$ , strictly submodular in  $n$  and  $a$ , and satisfies

<sup>6</sup>We thus also prove sufficiency for a class of models not covered by Jullien where the slope of the reservation utility satisfies a “shallow–steep” condition similar to stacking.

<sup>7</sup>Another example with sorting and incomplete information is Liu, Mailath, Postlewaite, and Samuelson (2014).

<sup>8</sup>We use “increasing” and “decreasing” in the weak sense, adding “strictly” when needed, and similarly with “positive” and “negative,” and “concave” and “convex.” We often write  $(f)_x$  for the total derivative of  $f$  with respect to  $x$ . We use the symbol  $=$ , to indicate that the objects on either side have strictly the same sign. We follow the hierarchy lemma, proposition, theorem. Wherever it is clear which firm we are talking about, we suppress the  $n$  superscript.

<sup>9</sup>Our model is equivalent to one with a single agent drawn from  $H$ . Log concavity is standard and in our setting avoids the need for ironing techniques.

<sup>10</sup>It is easy to include a type-independent or type-dependent outside option as well. See Section 5.3.



$\lim_{a \rightarrow \infty} c_a^n(a) = \infty$ . Here, higher indexed firms have lower marginal costs for producing quality.<sup>11</sup>

EXAMPLE 2—A labor market: Let  $\mathcal{V}^n(a)$  be the value to firm  $n$  of effort  $a$ , where  $\mathcal{V}^n$  is strictly supermodular, strictly increasing, and strictly concave with  $\lim_{a \rightarrow \infty} \mathcal{V}_a^n(a) = 0$ . Let the cost of effort to the worker be  $c(a) - a\theta$ , where  $c$  is convex with  $\lim_{a \rightarrow \infty} c_a(a) > 1 + (1/h(1))$ , so that  $\mathcal{V}(a, \theta) = \mathcal{U}(a) + a\theta = -c(a) + a\theta$ . Finally, assume that  $\mathcal{V}_a^n(0) - c_a(0) \geq 1/h(0)$ . All the assumptions hold if  $c(a) = 3a$  and  $\mathcal{V}^n(a) = \zeta^n + \beta^n \log(\rho + a)$ , where  $\rho > 0$  is sufficiently small and  $\beta^n$  is strictly increasing in  $n$ , so that higher indexed firms value effort more.

Let  $v_*^n(\theta) = \max_a (V^n(a) + a\theta)$  be the most surplus firm  $n$  can offer type  $\theta$  without losing money, and let  $\alpha_*^n(\theta)$  be the associated maximizer. We assume that each firm  $n$  is *relevant* in that there is  $\theta \in [0, 1]$  such that  $v_*^n(\theta) > \max_{n' \neq n} v_*^{n'}(\theta)$ . Relevance will be sufficient for all firms to be active in equilibrium. It says that firms with a lower marginal cost of quality have a sufficiently higher cost of providing low quality or that firms with a high marginal value for ability are comparatively poor at using less-able workers.<sup>12</sup> By relevance and strict supermodularity of  $V^n(a)$ , there are consecutive open intervals  $(a_e^{n-1}, a_e^n)$  of actions such that  $n$  is the most efficient firm at action  $a$ , where for  $1 \leq n < N$ ,  $V^n(a_e^n) = V^{n+1}(a_e^n)$ , and where  $a_e^0 = 0$  and  $a_e^N = \infty$ .<sup>13</sup>

Firms simultaneously offer menus of contracts, where firm  $n$ 's menu is a pair of functions  $(\alpha^n, t^n)$ , with  $\alpha^n(\theta)$  the action required of an agent who chooses firm  $n$  and announces type  $\theta$ , and  $t^n(\theta)$  the transfer to that agent. Contracts are exclusive: each agent can deal with only one firm. We rule out contracts that depend on other firms' offers.

Let  $v^n$  be the surplus function for an agent who takes the contract of firm  $n$ , given by

$$v^n(\theta) = \mathcal{V}(\alpha^n(\theta), \theta) - t^n(\theta) = \mathcal{U}(\alpha^n(\theta)) + \alpha^n(\theta)\theta - t^n(\theta).$$

It is without loss that firms offer incentive-compatible menus. Thus, going forward, we will describe menus by  $(\alpha^n, v^n)$ , where, as is standard, incentive compatibility is equivalent to requiring that the action schedule  $\alpha^n$  is increasing and (using that  $\mathcal{V}_\theta = a$ ) that  $v^n(\theta) = v^n(0) + \int_0^\theta \alpha^n(\tau) d\tau$  for all  $\theta$ , so that, in particular,  $v^n$  is convex, with derivative  $\alpha^n$  almost everywhere.<sup>14</sup> Since  $\alpha^n$  is increasing, it jumps at most a countable number of times, and so since  $h$  is atomless, it is without loss to assume  $\alpha^n$  at any  $\theta < 1$  to be right continuous and  $\alpha^n$  at 1 to be left continuous.

Firm  $n$ 's strategy set,  $S^n$ , is the set of incentive-compatible pairs  $s^n = (\alpha^n, v^n)$ . The joint strategy space is  $S = \times_n S^n$  with typical element  $s$ . Let  $s^{-n}$  be a typical strategy profile for firms other than  $n$ . Firm  $n$ 's profit on a type- $\theta$  agent who takes action  $a$  and is given utility  $v_0$  is  $\pi^n(\theta, a, v_0) = V^n(a) + a\theta - v_0$ . For any  $n$  and for any menu  $(\alpha, v)$  for  $n$ , we consolidate notation by writing  $\pi^n(\theta, \alpha, v)$  for  $\pi^n(\theta, \alpha(\theta), v(\theta))$ .

<sup>11</sup>The same example can be reinterpreted as a product market with quantity differentiation.

<sup>12</sup>For the parameterized labor market example above and for any given strictly increasing  $\{\beta^n\}$ , relevance is satisfied by an appropriate choice of  $\{\zeta^n\}$ . See Section 5.2 for more generality.

<sup>13</sup>Relevance holds if and only if for each firm  $n$ ,  $V^n$  is somewhere above the concave envelope of  $\max_{n' \neq n} V^{n'}$ , while  $(a_e^{n-1}, a_e^n)$  is nonempty if and only if  $V^n$  is somewhere above  $\max_{n' \neq n} V^{n'}$ . While relevance is sufficient for a firm to be active in equilibrium, we will see that  $(a_e^{n-1}, a_e^n)$  nonempty is necessary.

<sup>14</sup>See Armstrong and Vickers (2001) and Rochet and Stole (2002) for other examples of competition in utility.

After observing the posted menus, agents sort themselves to the most advantageous firm. Formally, for any  $n$  and  $s^{-n}$ , define the scalar-valued function  $v^{-n}$  given by  $v^{-n}(\theta) = \max_{n' \neq n} v^{n'}(\theta)$  as the most surplus offered by any of  $n$ 's competitors. As the maximum of convex functions,  $v^{-n}$  is convex. Let  $a^{-n}$  be the associated scalar-valued action function, so that  $a^{-n}$  is an increasing function almost everywhere equal to  $v_{\theta}^{-n}$ . From the point of view of firm  $n$ ,  $(a^{-n}, v^{-n})$  summarizes all the relevant information about the strategy profile of its opponents. Define  $\varphi^n(\theta, s)$  as the probability that  $n$  serves  $\theta$  given  $s$ . Thus,  $\varphi^n(\theta, s) = 0$  if  $v^n(\theta) < v^{-n}(\theta)$  and  $\varphi^n(\theta, s) = 1$  if  $v^n(\theta) > v^{-n}(\theta)$ . We assume ties are broken equiprobably.<sup>15</sup>

Define  $\Pi^n(s) = \int \pi^n(\theta, \alpha^n, v^n) \varphi^n(\theta, s) h(\theta) d\theta$  as the profit to firm  $n$  given strategy profile  $s$ . This reflects our assumptions that there are no capacity constraints and the technology is additively separable across agents. Because the optimal behavior of the agents is already embedded in  $\varphi$ , we can view the game as simply one among the firms, with strategy set  $S^n$  and payoff function  $\Pi^n$  for each  $n$ . Let  $BR^n(s) = \arg \max_{s^n \in S^n} \Pi^n(s^n, s^{-n})$ . Strategy profile  $s$  is a pure-strategy Nash equilibrium of  $(S^n, \Pi^n)_{n=1}^N$  if for each  $n$ ,  $s^n \in BR^n(s)$ .

#### 4. CHARACTERIZING EQUILIBRIUM

In this section, we begin with a set of necessary conditions that pure-strategy equilibria must satisfy. We show that a subset of these conditions is sufficient under a restriction, *stacking*, that the firms are sufficiently differentiated. Using this, we state an existence result under stacking. Finally, we show how our conditions allow straightforward numerical analysis of our model.

##### 4.1. Necessity

In this section, we study a set of necessary conditions for a Nash equilibrium in pure strategies under an equilibrium refinement. We will state our main theorem first, and then carefully define the conditions and explore their implications. But, as a rough guide, the equilibrium refinement, *no extraneous offers*, requires that the actions a given firm offers are continuous in type and constrained to a “reasonable” range. We show that in any equilibrium, each firm makes *positive profits* on each type, that no firm can profitably *poach* a (potentially distant) type from another firm by mimicking the offer of that firm to the type, and that there is *positive sorting* so that higher indexed firms serve higher sets of types. Finally, we show an *internal optimality* condition, pinning down the actions offered by a given firm to its interior types, and an *optimal boundaries* condition on the endpoints of the interval served by each firm, formally stated as follows.

**THEOREM 1—Necessity:** *Every pure-strategy Nash equilibrium with no extraneous offers has positive profits on each type served, no poaching, positive sorting, internal optimality, and optimal boundaries.*

As discussed, some of the “non-equilibrium” results in this section are related to **Julien (2000)**, but technical differences make it hard for us to import those results off the

---

<sup>15</sup> It can be shown that the set of pure-strategy Nash equilibria of our game is invariant if one replaces  $\varphi$  by any rule where for each  $\theta$ , all firms offering the highest surplus to  $\theta$  have a strictly positive probability of winning. This set is in turn weakly smaller than the set of equilibria with efficient tie-breaking. We conjecture that the sets are equal under the condition that no firm makes an offer that strictly loses money if accepted.



shelf. We provide intuition for those results here, and full and self-contained proofs in the Supplemental Material (Chade and Swinkels (2021)). Other properties are proven in Appendix A.

4.1.1. *Positive Profits*

The *positive profits* condition (PP) is satisfied if for each  $n$ , the probability that  $n$  serves an agent on whom he strictly loses money is 0. We prove—and use several times below—the stronger statement that for any  $s = (s^n, s^{-n})$  (equilibrium or not),  $s^n$  can be transformed to a strategy that is equivalent to  $s^n$  anywhere  $s^n$  earns positive profits, but eliminates any situation where  $s^n$  loses money. To see the intuition, let  $P$  be the set of types on which  $s^n$  makes money. Eliminate all menu items for types not in  $P$ . Types in  $P$  have fewer deviations available, and so incentive compatibility still holds. Types not in  $P$  who go to another firm save the firm money, and types not in  $P$  who now accept the same contract as a type in  $P$  are profitable, because, by private values, the firm is indifferent about the type of the agent who accepts an offer.<sup>16</sup>

By PP, there is no cross-subsidization: losing money on some types does not enhance the profits earned on others. Another key implication is that each firm earns strictly positive profits in equilibrium: since other firms do not lose money, and since  $(a_e^{n-1}, a_e^n)$  is nonempty, a firm that offers the menu  $(\alpha^n, v^n - \varepsilon)$  for  $\varepsilon$  sufficiently small will win a positive measure of agents and earn strictly positive profits on any agents served. See Proposition 2 in Appendix A.<sup>17</sup>

4.1.2. *No-Poaching*

Fix an equilibrium and let  $v^O(\cdot) = \max_n v^n(\cdot)$ , with  $O$  mnemonic for “oligopoly,” be the equilibrium surplus function. Let  $a^O$  be the associated action function, where, as before, we take  $a^O$  to be right continuous for  $\theta < 1$  and left continuous at 1. For any given  $a$ , let  $V^{(2)}(a)$  be the second largest element of  $\{V^n(a)\}_{n=1}^N$ . The *no-poaching condition* (NP) holds if for all  $\theta$ ,

$$v^O(\theta) \geq V^{(2)}(a^O(\theta)) + a^O(\theta)\theta,$$

so that  $\theta$  receives an amount at least equal to what the second most efficient firm could provide at the action implemented. That this holds only at the equilibrium action is important: it can be that a firm  $n'$  could profitably out-compete  $n$  on type  $\theta$  with another action, but does not do so, because it would attract some of its own types in a detrimental way.<sup>18</sup>

Moving  $v^O$  across the inequality, NP says that the second most efficient firm would lose money by poaching  $\theta$  at the current action. That is, if  $n$  is winning at  $\theta$ , then

$$\pi^n(\theta, \alpha^n, v^n) \leq V^n(\alpha^n(\theta)) - \max_{n' \neq n} V^{n'}(\alpha^n(\theta)).$$

This bound is strongest when firms have similar capabilities. It follows that under NP, each offer by firm  $n$  that is accepted in equilibrium is an element of  $[a_e^{n-1}, a_e^n]$ .

<sup>16</sup>See Jullien (2000, Lemma 3), or Supplemental Material, Proposition 5.

<sup>17</sup>Our proof that all firms make strictly positive profits uses in an essential way that any firm that matches the most favorable offer wins with strictly positive probability. If instead, indifferent types sort themselves to a firm that makes the most money on them, then a zero-profit equilibrium is that each firm offers the most surplus any firm can offer  $\theta$  without losing money. Such equilibrium can also be ruled out, as in Biglaiser and Mezzetti (1993), by directly assuming that firms do not make offers on which they strictly lose money if accepted.

<sup>18</sup>For PP and NP, the details of tie-breaking are inessential as long as  $\varphi^n$  is strictly positive wherever  $v^n(\theta) = v^{-n}(\theta)$ . Similarly, PP and NP do not use relevance or supermodularity of the firm’s payoff in  $n$  and  $a$ .

The intuition for the result is that if there is an interval of types where  $n$  is not winning always, but can make money by imitating the incumbent, then  $n$  can first mimic the behavior of the incumbent on those types and then add some small  $\varepsilon > 0$  in surplus everywhere, and so not affect the behavior of any type he is currently winning. As such, NP is about stealing the inframarginal agents of another firm.<sup>19</sup>

#### 4.1.3. *Positive Sorting*

Say that *quasi-positive sorting* (QPS) holds for strategy profile  $s$  if four things are true: First, for each firm  $n$ , there is a single nonempty interval  $(\theta_l^n, \theta_h^n)$  of agents such that firm  $n$  serves a full measure of the agents in that interval (but may be serving some zero-measure set of types each with probability less than 1). Second, these intervals are ordered, so that  $\theta_h^n \leq \theta_l^{n+1}$  for all  $n$ . Third,  $\theta_l^1 = 0$ , and  $\theta_h^N = 1$ . Finally, if  $\theta_h^n < \theta_l^{n+1}$ , then for each type  $\theta \in (\theta_h^n, \theta_l^{n+1})$ , both firms are offering action  $a_e^n$  and transferring all surplus,  $V^n(a_e^n) + a_e^n \theta$ , to the agent, so that each firm is winning half the time and profits are zero on these types.

We assert that any pure-strategy equilibrium has QPS. To see the intuition, fix  $\theta' > \theta$ . By incentive compatibility, the equilibrium action of  $\theta'$  is at least as high as that of  $\theta$ . But  $V^n$  is strictly supermodular in  $n$  and  $a$ . Hence, if  $n$  sometimes serves  $\theta'$  and  $n' > n$  sometimes serves  $\theta$ , then, by PP and NP, either  $n$  will want to always serve  $\theta$  or  $n'$  will want to always serve  $\theta'$ , a contradiction. The only exception is if both firms are indifferent about hiring both  $\theta$  and  $\theta'$ , and this can only happen if actions are constant and equal to  $a_e^n$  on the tied interval, and profits are dissipated. See Appendix A for the proof.

Say that an equilibrium has *positive sorting* (PS) if on  $(\theta_l^n, \theta_h^n)$ , firm  $n$  serves each type with probability 1. Say that  $s$  has *strictly positive sorting* (SPS) if in addition  $\theta_h^n = \theta_l^{n+1}$  for all  $1 \leq n < N$ , so that there are no intervals of ties. Under SPS, there will typically be gaps in the action level as one moves from one firm to the next. Figure 1 shows a typical example with SPS and four firms. We explain how the figure was generated in Section 4.3.

The conditions QPS and PS differ when an outside firm makes offers that set a floor on the surplus offered by the active firm but bind only at a zero-measure set of types. Such offers are not without economic rationale. Indeed, in a contestable market (see, for example, Baumol (1988)), a firm may be inactive in equilibrium, but still constrain the active firms. Here, such a thing could occur on a type-by-type basis.<sup>20</sup>

#### 4.1.4. *No Extraneous Offers*

Let us now introduce an equilibrium refinement that allows us to focus on settings where PS holds. By NP, we know that with probability 1, the actions of firm  $n$  that are accepted in equilibrium are elements of  $[a_e^{n-1}, a_e^n]$ , the set where  $n$  is the most efficient. In Lemma 4 (Appendix A), we also show that any best response for  $n$  must be continuous on  $(\theta_h^{n-1}, \theta_l^{n+1})$ , the region over which  $n$  ever wins. Say that an equilibrium has *no extraneous offers* (NEO) if each function  $\alpha^n$  is continuous everywhere and takes on values that fall in  $[a_e^{n-1}, a_e^n]$ .<sup>21</sup> Appendix A shows that under NEO, any equilibrium has PS.

<sup>19</sup>See Appendix A, Proposition 3 for a proof. While Jullien (2000, Lemma 2) is somewhat related, our endogenous outside option introduces an important and economically interesting new dimension to the analysis.

<sup>20</sup>Indeed, without some refinement, there are equilibria in which firms make offers that would lose money if accepted, but are accepted by a zero-measure set of types and so do not hurt the firm making them.

<sup>21</sup>Note well that this is an equilibrium refinement, not a restriction to the strategy spaces. At the heart of the proof of NP is that when a firm is too greedy, other firms can imitate it.

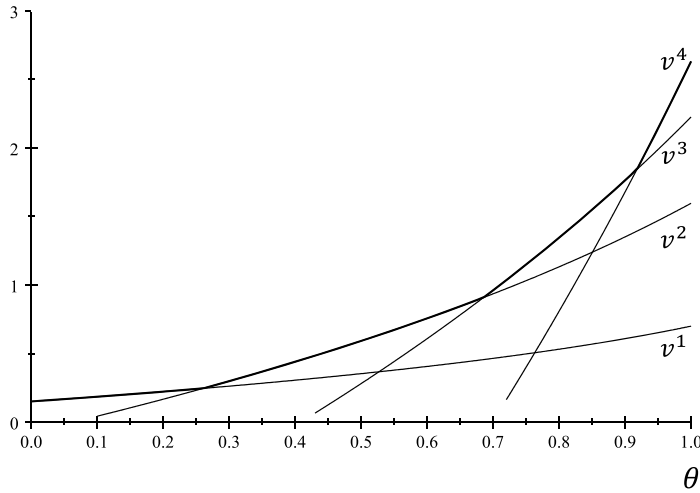


FIGURE 1.—An equilibrium with four firms. The curves  $v^n$  are the equilibrium surplus functions. Each firm serves the types where its surplus function is highest. The agent receives surplus indicated by the thick locus. Upward kinks in this locus reflect jumps in the action.

4.1.5. *Internal Optimality*

Each firm will distort the action schedule so as to reduce information rents on its interior types. Fix  $n$  and for  $\kappa \in [0, 1]$ , define  $\gamma^n$  by

$$\pi_a^n(\theta, \gamma^n(\theta, \kappa), v^n(\theta)) = \frac{\kappa - H(\theta)}{h(\theta)}, \tag{1}$$

recalling that  $\pi_a^n$  does not depend on  $v^n$ , so that this is well defined. Strategy profile  $s$  satisfies *internal optimality* (IO) if for each  $n$ , there is  $\kappa^n \in [H(\theta_l^n), H(\theta_h^n)]$ , where  $\kappa^1 = 0$  and  $\kappa^N = 1$ , such that  $\alpha^n(\cdot) = \gamma^n(\cdot, \kappa^n)$  on  $[\theta_l^n, \theta_h^n]$ .<sup>22,23</sup> By IO, there is a type  $\theta_0^n(\kappa^n) \in [\theta_l^n, \theta_h^n]$  satisfying  $H(\theta_0^n(\kappa^n)) = \kappa^n$ . To simplify the notation, henceforth we will omit the argument  $\kappa$  from  $\theta_0^n$ . Actions are distorted down below  $\theta_0^n$ , distorted up above  $\theta_0^n$ , and are efficient at  $\theta_0^n$ . Since  $\gamma^n(\cdot, \kappa^n)$  is strictly increasing, an economic implication of IO is that there is complete sorting within the interval of types uniquely served by each firm  $n$ .

We will shortly relate  $\kappa$  to a Lagrange multiplier in a suitable problem. But to see intuition for IO, note that since  $\kappa^N = 1$ , (1) reduces for firm  $N$  to the standard equation (Mussa and Rosen (1978), Maskin and Riley (1984)) for a monopolist screening an agent of unknown type. To reduce information rents while retaining the lowest type served, firm  $N$  lowers the slope of the surplus function by distorting actions downward. In contrast, for firm 1, where the only participation constraint that binds is for the top type served, distorting actions upward steepens the surplus function, reducing information rents. Indeed,

<sup>22</sup>Since  $V_a^n(0) \geq 1/h(0)$  and  $\lim_{a \rightarrow \infty} V_a^n(a) < -1 - (1/h(1))$ , (1) has an interior solution. Log concavity of  $H$  and  $1 - H$  imply that  $(\kappa - H(\cdot))/h(\cdot)$  is decreasing for all  $\kappa \in [0, 1]$ . To see this, differentiate to obtain  $((\kappa - H)/h)_\theta = -1 - (\kappa - H)h'/h^2$ , where  $h'(\theta) \leq 0$ , and since  $1 - H$  is strictly log-concave,  $-1 - (\kappa - H)h'/h^2 \leq -1 - (1 - H)h'/h^2 = ((1 - H)/h)_\theta$ . If  $h'(\theta) > 0$ , then the result follows since  $H$  is strictly log-concave. Thus, since  $\pi_{a\theta} = 1 > 0$ , we obtain that  $\gamma^n$  is strictly increasing in  $\theta$ . Similarly,  $\gamma^n$  is strictly decreasing in  $\kappa$ .

<sup>23</sup>See Jullien (2000, Theorem 1 and Proposition 2) for a similar result, and the Supplemental Material for a self-contained proof that deals with the details of our environment.

$\kappa^1 = 0$  and so  $\pi_a^1$  is everywhere negative. For intermediate firms, the action is distorted down below  $\theta_0^n$  but up for higher types. This maintains the surplus on the boundary types, but lowers the information rents of interior types.

To see the functional form of  $\gamma^n$ , and build structure that we will need when we turn to sufficiency and existence, fix boundary points  $\theta_l$  and  $\theta_h$  for firm  $n$ , and let  $\mathcal{P}(\theta_l, \theta_h)$  be the following problem for firm  $n$  (per our convention, we omit the superscript  $n$  for simplicity):

$$\max_{(\alpha, v)} \int_{\theta_l}^{\theta_h} \pi(\theta, \alpha, v) h(\theta) d\theta$$

$$\text{s.t. } v(\theta_l) \geq v^{-n}(\theta_l), \tag{2}$$

$$v(\theta_h) \geq v^{-n}(\theta_h), \tag{3}$$

$$v(\theta) = v(0) + \int_0^\theta \alpha(\tau) d\tau \quad \text{for all } \theta. \tag{4}$$

This relaxes  $n$ 's problem, as we drop monotonicity of  $\alpha$ , ignore the outside option except at  $\theta_l$  and  $\theta_h$ , and relax the constraint at  $\theta_l$  and  $\theta_h$ . Let  $\iota(\theta_l, \theta_h, \kappa) = v^{-n}(\theta_h) - v^{-n}(\theta_l) - \int_{\theta_l}^{\theta_h} \gamma(\theta, \kappa) d\theta$  and let  $\tilde{\kappa}(\theta_l, \theta_h) = \arg \min_{\kappa \in [H(\theta_l), H(\theta_h)]} |\iota(\theta_l, \theta_h, \kappa)|$ . That is, subject to  $\kappa$  lying in  $[H(\theta_l), H(\theta_h)]$ , the firm comes as close as possible to matching the rise  $\int_{\theta_l}^{\theta_h} \gamma(\theta, \kappa) d\theta$  in the surplus it offers to the increase  $v^{-n}(\theta_h) - v^{-n}(\theta_l)$  in the outside option.<sup>24</sup> We have the following lemma.

**LEMMA 1—Relaxed Problem:** *Problem  $\mathcal{P}(\theta_l, \theta_h)$  has a solution  $\tilde{s}(\theta_l, \theta_h) = (\tilde{\alpha}, \tilde{v})$ . On  $(\theta_l, \theta_h)$ ,  $\tilde{\alpha}$  is uniquely defined and equal to  $\gamma(\cdot, \tilde{\kappa}(\theta_l, \theta_h))$ .<sup>25</sup> If  $\tilde{\kappa}(\theta_l, \theta_h) > H(\theta_l)$ , then  $\tilde{v}(\theta_l) = v^{-n}(\theta_l)$ , and if  $\tilde{\kappa}(\theta_l, \theta_h) < H(\theta_h)$ , then  $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$ .*

To interpret this result, let  $\eta$  be the shadow value of increasing the surplus of type  $\theta_h$  holding fixed the surplus of type  $\theta_l$ . One way firm  $n$  can achieve this increase is to raise the action at any given interior type  $\theta$ . This has benefit  $\pi_a(\theta, \tilde{\alpha}, \tilde{v})$  on the  $h(\theta)$  types near  $\theta$ , and raises the surplus on the  $H(\theta_h) - H(\theta)$  types between  $\theta$  and  $\theta_h$ , and so  $\eta = -\pi_a(\theta, \tilde{\alpha}, \tilde{v})h(\theta) + H(\theta_h) - H(\theta)$ . At an optimum, this expression must hold for all types, since otherwise the firm could profitably raise the action at one  $\theta$  and lower it at another, leaving  $v(\theta_h)$  unaffected. In particular, since  $\pi_a = 0$  at  $\theta_0$  and since  $H(\theta_0) = \kappa$ , we have  $\eta = H(\theta_h) - \kappa$ . Substituting and rearranging yields (1).<sup>26</sup> To use this result to show IO, we show that if  $\alpha \neq \gamma(\cdot, \tilde{\kappa})$ , then we can perturb  $\alpha$  in the “direction” of  $\gamma(\cdot, \tilde{\kappa})$  strictly profitably.<sup>27</sup>

<sup>24</sup>By footnote 22,  $\gamma_\kappa < 0$ , and so  $\iota_\kappa > 0$ . Thus,  $\tilde{\kappa}$  is  $H(\theta_l)$  if  $\iota(\theta_l, \theta_h, H(\theta_l)) > 0$ , is  $H(\theta_h)$  if  $\iota(\theta_l, \theta_h, H(\theta_h)) < 0$ , and is the solution to  $\iota(\theta_l, \theta_h, \kappa) = 0$  otherwise, and hence  $\tilde{\kappa}$  is well defined and continuous.

<sup>25</sup>To make  $\tilde{s}$  uniquely defined everywhere, define  $\tilde{\alpha}(\theta) = \tilde{\alpha}(\theta_h)$  for  $\theta \geq \theta_h$  and  $\tilde{\alpha}(\theta) = \tilde{\alpha}(\theta_l)$  for  $\theta \leq \theta_l$ .

<sup>26</sup>For further intuition for  $\kappa$ , assume that (2) binds at the optimum (a similar argument holds if (3) binds, and we show in the proof of the lemma that at least one of them must bind). Then by the envelope theorem, one can verify that a marginal increase in  $v^{-n}(\theta_h)$  reduces profits by  $H(\theta_h) - \kappa$ , and that a marginal increase in  $v^{-n}(\theta_l)$  reduces profits by  $\kappa - H(\theta_l)$ . That is,  $\kappa$  pins down the shadow values of constraints (2) and (3).

<sup>27</sup>See the Supplemental Material for a proof of the lemma and a proof (Proposition 6) that this implies IO.

4.1.6. *Optimal Boundaries*

Strategy profile  $s$  satisfies the *optimal boundary condition* (OB) if for  $\theta = \theta_l^n$  and  $\theta = \theta_h^n$ ,

$$\pi^n(\theta, \alpha^n, v^n) + \pi_a^n(\theta, \alpha^n, v^n)(a^{-n}(\theta) - \alpha^n(\theta)) = 0, \tag{5}$$

where we discard the condition at  $\theta_l^1 = 0$  and at  $\theta_h^N = 1$ .<sup>28</sup>

Under OB, small changes in the interval of served types do not pay. It contrasts with NP, which is about stealing potentially distant agents. To see the intuition for OB, fix  $n$  and increase the action of types near  $\theta_h$  a little. This has direct benefit  $\pi_a(\theta_h, \alpha, v)h(\theta_h)$ , but raises  $v(\theta_h)$ . As  $v(\theta_h)$  is raised,  $\theta_h$  increases at rate  $1/(a^{-n}(\theta_h) - \alpha(\theta_h))$  since  $\alpha(\theta_h)$  is the slope of  $v$  at  $\theta_h$  and  $a^{-n}(\theta_h)$  is the slope of  $v^{-n}$ . Hence, the profits from the new types served are  $\pi(\theta_h, \alpha, v)h(\theta_h)/(a^{-n}(\theta_h) - \alpha(\theta_h))$ . Cancelling  $h(\theta_h)$  and rearranging yields (5), and similarly for  $\theta_l$ .

We will use our next simple lemma repeatedly. The slope of profit with respect to  $\theta$  has the sign of  $\pi_a\alpha_\theta$ , and if the action profile is of the form given by (1), then profits are strictly single-peaked.

LEMMA 2—Profit Single-Peaked: *For any  $(\alpha, v) \in S^n$ ,*

$$\frac{d}{d\theta}\pi(\theta, \alpha, v) = \pi_a(\theta, \alpha, v)\alpha_\theta(\theta). \tag{6}$$

*If  $\alpha = \gamma(\cdot, H(\theta_0))$ , then  $\pi(\cdot, \alpha, v)$  is strictly single-peaked with peak at  $\theta_0$ .*

To see (6), note that by definition of  $\pi$ ,  $\frac{d}{d\theta}\pi(\theta, \alpha, v) = \pi_\theta(\theta, \alpha, v) + \pi_a(\theta, \alpha, v)\alpha_\theta(\theta) - v_\theta(\theta)$  and that  $\pi_\theta(\theta, \alpha, v) = \alpha(\theta) = v_\theta(\theta)$ . If  $\alpha = \gamma(\cdot, H(\theta_0))$ , then from footnote 22,  $\alpha_\theta > 0$ , and  $\pi_a(\theta, \alpha, v)$  has strictly the same sign as  $\theta_0 - \theta$ . Hence,  $\pi$  is strictly single-peaked at  $\theta_0$ .

That profits are strictly single-peaked at  $\theta_0$  has some intuition: For intermediate firms, customers in the middle of the participation range find neither of the alternative firms very attractive, and so are the easiest from whom to extract rents. Similarly, for the end firms, it is the extreme types from whom it is easiest to extract rents.

Let us now show that  $\kappa$  is interior for  $n \notin \{1, N\}$ . If  $\kappa = H(\theta_h)$ , then by Lemma 2,  $\pi(\cdot, \alpha, v)$  is strictly increasing on  $(\theta_l, \theta_h)$  and so, since  $\pi(\theta_l, \alpha, v) \geq 0$ , it follows that  $\pi(\theta_h, \alpha, v) > 0$ . But since  $\kappa = H(\theta_h)$ , we also have  $\pi_a(\theta_h, \alpha, v) = 0$ , and so (5) is violated. Essentially, if  $\kappa = H(\theta_h)$ , then increasing the action on types near  $\theta_h$  has second-order efficiency costs but gains some extra agents on whom profits are strictly positive. Similarly,  $\kappa > H(\theta_l)$ .

An important implication of Lemma 2 is that, in equilibrium,  $\pi$  is strictly positive on  $(\theta_l, \theta_h)$ . This follows since by PP,  $\pi$  is positive at  $\theta_l$  and  $\theta_h$ , and since  $\alpha$  is given by (1) on  $[\theta_l, \theta_h]$ , and so  $\pi$  is strictly single-peaked on  $[\theta_l, \theta_h]$ . For the boundary types  $\theta_l$  and  $\theta_h$ , if there is a region of overlap between the two firms, profits on these types are zero. Consider the case depicted in Figure 1, where the surplus functions cross strictly and so the implemented action jumps at the boundary. Then, since we have already argued that  $\theta_0 < \theta_h$  for  $n < N$ , the term  $\pi_a(\theta_h, \alpha, v)(a^{-n}(\theta_h) - \alpha(\theta_h))$  in (5) is strictly negative. Thus,  $\pi(\theta_h, \alpha, v)$  is strictly positive, and similarly for  $\pi(\theta_l, \alpha, v)$ . The difference in their technologies implies that neither firm can profitably imitate the other despite strictly positive profits.

<sup>28</sup>A close relative is Jullien (2000, Theorem 2). A proof that takes care of the possibility that  $a^{-n} = \alpha^n$  at either boundary is provided in the Supplemental Material.

4.2. *Sufficiency and Existence*

We begin by making an assumption that eliminates ties at the boundaries between firms.

DEFINITION 1: Stacking is satisfied if for all  $n < N$ ,  $\gamma^{n+1}(\cdot, 1) > \gamma^n(\cdot, 0)$ .

Under stacking, firm  $n + 1$ 's action schedule lies strictly above that of firm  $n$ , and so surplus functions cross strictly. Stacking holds if firms are sufficiently differentiated (see, for example, the numerical example in the Supplemental Material, Section S2).<sup>29</sup> For given  $n$  and  $s^n$ , say that  $s^n$  and  $\hat{s}^n$  are *equivalent* if  $s^n$  and  $\hat{s}^n$  differ only where neither ever wins. Two strategy profiles are equivalent if they are equivalent for each  $n$ .

THEOREM 2—Sufficiency and Existence: *Assume stacking. Then any strategy profile satisfying PS, IO, and OB is equivalent to a Nash equilibrium, and a Nash equilibrium exists.*

Crucially, under stacking, the nonlocal condition NP can be dropped, leaving only local conditions. We defer discussion of the (surprisingly intricate) proof to Section 6.

4.3. *Numeric Analysis*

Theorem 2 facilitates numeric analysis. By PS, there are  $N - 1$  interior boundary points  $\theta^n$  between consecutive firms, and using IO, each firm's behavior is characterized by a "slope"  $\kappa^n$  and an intercept  $v^n(0)$ . Since  $\kappa^1$  and  $\kappa^N$  are fixed, we have  $3N - 3$  unknowns. But at each interior boundary point, each relevant firm has to satisfy OB, and the surpluses offered by the firms must agree, for a total of  $3N - 3$  equations. By sufficiency, this set of equations characterizes an equilibrium, and so by existence, it has a solution. Finding these solutions numerically is trivial. Figure 1 carries out this process for four firms with  $\mathcal{V}^n(a) = \zeta^n + \beta^n \log a$  and agents with  $\mathcal{V}(a, \theta) = -(3 - \theta)a$ . It assumes that  $h$  is uniform and that  $\beta^1 = 1, \beta^2 = 4, \beta^3 = 9, \beta^4 = 20, \zeta^1 = 2.5, \zeta^2 = 3, \zeta^3 = -2$ , and  $\zeta^4 = -23$ . See the Supplemental Material, Section S2 for further details. We repeatedly extend this example going forward.

5. IMPLICATIONS AND APPLICATIONS

5.1. *Who Does Incomplete Information Help or Hurt?*

Consider a version of our model with complete information. A monopolist is better off, since, compared to incomplete information, it can undo any inefficiency and then extract all the surplus, leaving all types worse off. In oligopoly, there is another effect: competition under complete information increases the agents' outside option relative to incomplete information. The reason for this is that with incomplete information, an offer that both attracts and earns profits on some new types for firm  $n$  might not be made because it would also attract some of  $n$ 's existing types at lower profits. With complete information, there is no such trade-off. Indeed, in equilibrium there is positive sorting and each type is served efficiently and receives surplus equal to what the second most efficient firm can provide.<sup>30</sup>

<sup>29</sup>If firms are not very differentiated, then equilibria must involve ties. To see this, let  $N = 2$  and  $\gamma^2(\cdot, 1) < \gamma^1(\cdot, 0)$ . If there are no ties, then  $\theta_1^2 = \theta_1^1$ , and so  $\alpha^1(\theta_1^1) = \gamma^1(\theta_1^1, 0) > \gamma^2(\theta_1^1, 1) = \alpha^2(\theta_1^1)$ , contradicting PS.

<sup>30</sup>To avoid an uninteresting openness issue, we assume for the complete information case that, as in Biglaiser and Mezzetti (1993), agents break ties in favor of the firm that earns more profit in serving them, and also that no firm makes an offer that they would lose money on if accepted.



When comparing pure-strategy equilibria under complete and under incomplete information, a simple structure arises. For  $1 \leq n < N$ , let  $\theta_*^n$  be the boundary point between the intervals of types where  $n$  and  $n + 1$  are the most efficient firms to serve  $\theta$ ; that is,  $v_*^n(\theta_*^n) = v_*^{n+1}(\theta_*^n)$ . Such a point exists by relevance and is unique by strict supermodularity of  $V^n(a)$ . Let  $\theta^n$  be the boundary point between  $n$  and  $n + 1$  in the equilibrium under incomplete information.

THEOREM 3—Welfare: *Under stacking, the following statements hold:*

- (i) *For each  $1 \leq n \leq N - 1$ , an interval of types containing  $\theta_*^n$  and  $\theta^n$  is strictly better off under complete information.*
- (ii) *If any type served by firm  $n$  under incomplete information is strictly better off than under complete information, then there is a single subinterval of  $[\theta_*^{n-1}, \theta_*^n]$  of such types.*
- (iii) *All types may be strictly better off under complete information.*

The proof (Appendix B) also shows that there is an open interval containing  $\theta_*^n$  where the firm is harmed under complete information. For intuition, with complete information,  $\theta_*^n$  earns the efficient surplus and so is better off than with incomplete information, while the firms earn zero on  $\theta_*^n$  and so are worse off. *Contrary* to monopoly, some types—perhaps *all*—are benefited by complete information. See Figure 2, which builds on Fig-

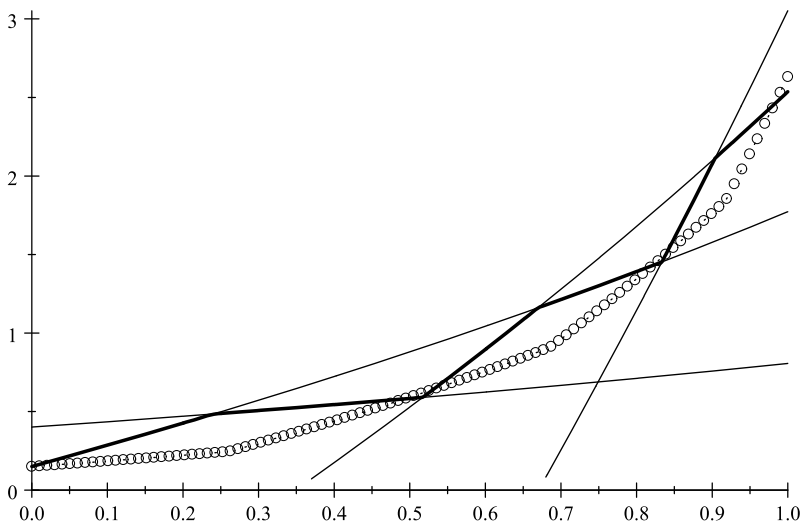


FIGURE 2.—Complete versus incomplete information. The thin lines are the efficient surplus each firm can provide,  $v_*^n$ , in the setting of Figure 1. The thick line is the amount that the agent receives under complete information. The bubbled line is the surplus in the incomplete information oligopoly. Complete information is preferred by intervals that include any type where two firms can provide the efficient surplus (at downward kinks of the thin line) and on the boundary between two types in equilibrium (at upward kinks of the solid line). In this example, incomplete information is preferred by types on a right neighborhood of  $\theta = 0$ , on an interval around the point where  $v_*^1 = v_*^3$  and similarly where  $v_*^2 = v_*^4$ , and on a left neighborhood of  $\theta = 1$ .

ure 1.<sup>31</sup> Such examples exist for any  $N$ , although we show next that in both cases, surplus converges to the efficient level as  $N$  grows.

### 5.2. The Competitive Limit

Consider now a setting where firms can enter at a fixed cost  $F > 0$  and then choose freely from a set of potential technologies parameterized by  $z \in [0, \bar{z}]$ . With some abuse of notation, let  $V(a, z) + a\theta$  be the match surplus from action  $a$ , technology  $z$ , and type  $\theta$ , where  $V$  is  $C^2$ , strictly concave with Hessian with strictly positive determinant, and strictly supermodular. To avoid boundary cases, we assume that for each  $a$ ,  $V(a, \cdot)$  has an interior maximum, and, similar to before, that  $V_a(0, 0) > 1/h(0)$  and  $\lim_{a \rightarrow \infty} V_a(a, \bar{z}) < -1 - (1/h(1))$ . Let  $z_l$  and  $z_h$  be the technologies that are best suited to serve types 0 and 1, respectively.<sup>32</sup>

In a *pure-strategy equilibrium with endogenous entry*, the  $N$  extant firms each earn at least  $F$ , but no new entrant can do so. Note that the most type  $\theta$  could possibly hope for is  $v_*(\theta) = \max_{a,z}(V(a, z) + a\theta)$ . We have the following theorem, which builds only on the necessary conditions we have derived for a pure-strategy equilibrium.<sup>33</sup>

**THEOREM 4—Limit Efficiency:** *There is  $\rho \in (0, \infty)$  such that in any pure-strategy equilibrium with endogenous entry and NEO,  $1/(\rho F^{1/3}) \leq N \leq (\rho/F^{1/3}) + 2$ . The profit per type,  $\pi$ , and the difference between what each type  $\theta$  earns and  $v_*(\theta)$  are each of order  $1/N^2$ .*

The heart of the proof is to show that as  $F$  goes to zero, the number of firms grows like  $1/F^{1/3}$ , these firms are located without large gaps, actions are efficient, and all surplus goes to the agents. The theorem implies that industry profits converge to zero like  $1/N^2$ , as does the total expenditure on entry costs,  $N \times F$ .

Intuitively, any given gap between firms between  $z_l$  and  $z_h$  implies that some consumers are being poorly served, which leads to a profitable entry opportunity when  $F$  is small. But when the firms are tightly packed, they each serve a small interval of types, and since the action is efficient for some interior type of the firm, inefficiencies in action choices are also small. And for each type, there is another firm that is nearly as efficient at the induced action, and so the gains in the market go to the agent. See the Supplemental Material, Section S4.1 for details.

### 5.3. Multi-Technology Monopoly and Oligopoly With a Merger

Assume that a single firm  $M$  controls more than one technology  $V^n$ . For example, LVMH, through a sequence of mergers and acquisitions, controls a set of technologies (brands with different production facilities) specialized to different quality points in multiple luxury segments. How does the presence of firm  $M$  affect the market? We approach this question through a sequence of steps, each of independent interest. Omitted proofs can be found in Appendix B.

<sup>31</sup>See the Supplemental Material for an example where  $h$  is changed so that all types strictly prefer complete information. Such examples are also easy to build with two firms when  $\mathcal{V}^n$  is linear and effort is constrained to an interval.

<sup>32</sup>See the Supplemental Material, Section S4.1, Step 0, for details. We show there that for any finite  $N$  and for  $z_l \leq z^1 < \dots < z^N \leq z_h$ , we can take  $V^n(a) = V(a, z^n)$  and have a market satisfying the conditions of Section 3. For example, let  $V(a, z) = a - a^2/2 - (a - z)^2$ , take  $[z_l, z_h] = [1, 2]$ , and take  $V^n(a) = V(a, 1 + n/N)$ .

<sup>33</sup>As an important caveat, recall that we do not have a pure-strategy existence result when firms are minimally differentiated, as they will be for large  $N$ . We leave for further work such an existence result or, alternatively, a result showing limit efficiency when strategies are mixed.

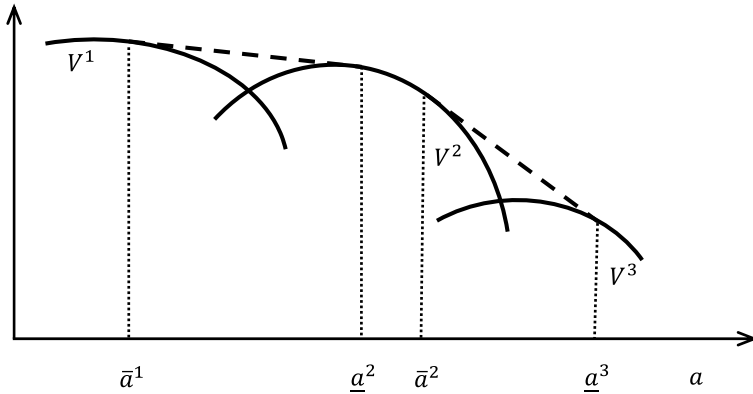


FIGURE 3.—A multi-technology firm. The firm controls three technologies,  $V^1$ ,  $V^2$ , and  $V^3$ . Below  $\bar{a}^1$   $\bar{V} = V^1$ , on  $(\underline{a}^2, \bar{a}^2)$ , we have  $\bar{V} = V^2$ , and above  $\underline{a}^3$ , we have  $\bar{V} = V^3$ . The dotted lines complete the concave envelope of the technologies.

5.3.1. *A Multi-Technology Monopoly*

Let a single firm  $M$  control technologies  $n_1, \dots, n_h$ , where each technology has the properties we have previously assumed hold for the technology of an individual firm and where each technology is relevant in the sense already defined. We assume for simplicity that the setting is synergy-free, in that for any given type, firm  $M$  can choose which technology to use, but cannot combine aspects of the different technologies into a hybrid technology.<sup>34</sup>

Assume first that  $M$  is a monopolist facing a convex outside option  $\bar{u}$  which, per Section 4.2, is first below  $\gamma^{n_1}(\cdot, 1)$  and then above  $\gamma^{n_h}(\cdot, 0)$ .<sup>35</sup> To analyze this problem, let  $\bar{V}$  be the concave envelope of  $\max\{V^{n_1}, \dots, V^{n_h}\}$  (see Figure 3). By relevance, for each  $n \in \{n_1, \dots, n_h\}$ ,  $\bar{V}$  equals  $V^n$  over some strictly positive-length interval  $[\underline{a}^n, \bar{a}^n]$ , where these intervals are disjoint, and where  $\underline{a}^{n_1} = 0$  and  $\bar{a}^{n_h} = \infty$ . We will show that  $M$  acts as if it had technology  $\bar{V}$ , and so we can apply all of what we already know about a single firm. Since  $\bar{V}$  will have a linear segment as it moves from each interval  $[\underline{a}^n, \bar{a}^n]$  to the next, modify the action schedule to choose the largest action consistent with IO; that is,

$$\gamma^M(\theta, \kappa) = \max \left\{ a \left| \bar{V}_a(a) + \theta = \frac{\kappa - H(\theta)}{h(\theta)} \right. \right\}, \tag{7}$$

noting that  $\bar{V}_a(a) + \theta$  plays the role of  $\pi_a$ . Where  $\gamma^M(\theta, \kappa) \in (\underline{a}^n, \bar{a}^n)$ , we have that  $\bar{V}$  is strictly concave, and so  $\gamma^M(\cdot, \kappa) = \gamma^n(\cdot, \kappa)$  and  $\bar{V} = V^n$ . Where production moves from technology  $n$  to  $n + 1$ , the action schedule  $\gamma^M(\cdot, \kappa)$  jumps from  $\bar{a}^n$  to  $\underline{a}^{n+1}$ , and we (arbitrarily) chose  $\underline{a}^{n+1}$  at such points.<sup>36</sup> Again,  $\gamma^M(\cdot, \kappa) = \gamma^n(\cdot, \kappa)$  and  $\bar{V} = V^n$ .

<sup>34</sup>So, in our setting, a data-provision firm which owns two technologies, one of which has low installation costs but a high marginal cost of bandwidth, and a second with high installation costs but a low cost of bandwidth, can choose which one to use to serve any given customer, but cannot combine them to create a technology with low installation costs and low bandwidth costs. Our analysis sets an important baseline even if such synergies exist.

<sup>35</sup>With a more general  $\bar{u}$ , firm  $M$  might choose to serve several intervals of types, complicating the analysis.

<sup>36</sup>The choice of action at this finite set of points is irrelevant to the surplus integrals.

With this modification, our previous analysis goes through. A firm with technology  $\bar{V}$  optimally operates on an interval  $[\theta_l^M, \theta_h^M]$ , there is a single  $\kappa \in [H(\theta_l^M), H(\theta_h^M)]$  such that all active technologies operate according to  $\gamma^M(\cdot, \kappa)$ , and the optimal boundary condition for the highest and lowest types served is the same as in our previous analysis.<sup>37</sup> To see that this solution is also optimal for  $M$  (which has technology  $\max_{n \in \{n_l, \dots, n_h\}} V^n$  rather than  $\bar{V}$ ), note that  $\bar{V}$  is at least as big as  $\max_{n \in \{n_l, \dots, n_h\}} V^n$  for all  $a$ , but the two are equal everywhere in the range of  $\gamma^M$ . Thus, the solution is feasible and, hence, optimal for the merged firm.

Let us turn to the optimal boundaries between  $M$ 's constituent operating technologies. For any given  $\theta$  and  $\kappa$ , let  $n(\theta, \kappa)$  be the unique technology for which  $\gamma^n(\theta, \kappa) \in [\underline{a}^n, \bar{a}^n]$ . From (7),  $n(\cdot, \kappa)$  does not depend on the outside option schedule  $\bar{u}$ , so if  $n$  and  $n + 1$  remain active, then the boundary point between them,  $\theta^{M,n}$ , depends only on  $\kappa$ . Also from (7), for each  $\theta$ ,  $\gamma^M(\theta, \kappa)$  is a maximizer of  $\bar{V}(a) + \theta a + ((\kappa - H(\theta))/h(\theta))a$ . Thus, since both  $\underline{a}^{n+1}$  and  $\bar{a}^n$  are maximizers at  $\theta^{M,n}$ , we can rearrange to arrive at

$$\pi^n(\theta^{M,n}) - \pi^{n+1}(\theta^{M,n}) + \frac{\kappa - H(\theta^{M,n})}{h(\theta^{M,n})}(\underline{a}^{n+1} - \bar{a}^n) = 0. \tag{8}$$

This differs from OB by adding the term  $-\pi^{n+1}(\theta^{M,n})$ , reflecting that the now-internalized externality that the customer gained for  $n$  is lost by  $n + 1$ .

### 5.3.2. Oligopoly versus Monopoly With a Fixed Market Size

Next let us compare the outcome of the monopolist firm  $M$  with a setting where firms  $n_l, \dots, n_h$ , with the associated technologies, compete oligopolistically given a convex outside option  $\bar{u}$  (which is again assumed to be first below  $\gamma^{n_l}(\cdot, 1)$  and then above  $\gamma^{n_h}(\cdot, 0)$ ). In this subsection, we assume that  $M$  is forced to serve the same aggregate set of types as do the constituent firms in oligopoly, but can adjust each type's action and the allocation of types across its constituent firms. Interpreting  $M$  as the result of a merger among the constituent firms and noting that such "must-serve" conditions are often imposed by antitrust regulators as part of a merger approval, this setting is of economic interest. It will also illuminate the conflicting forces when we deal with a merger in an oligopoly setting.

We now show that  $M$  offers less surplus to every interior type. Thus, to protect consumers or workers after a merger, it is not enough to require the merged firm to serve the same set of types, since it will reoptimize its rent extraction so as to hurt them all.

**THEOREM 5—Fixed Span:** *Let  $[\theta_l, \theta_h]$  be the set of types served in oligopoly by firms  $n_l, \dots, n_h$  facing outside option  $\bar{u}$ . If forced to serve exactly  $[\theta_l, \theta_h]$ , then firm  $M$  will choose  $\kappa$  in  $(\kappa^{n_l}, \kappa^{n_h})$ . All types in  $(\theta_l, \theta_h)$  are strictly worse off, with an interval of low types receiving a strictly lower action than before and an interval of high types receiving a strictly higher action than before.*

Intuitively, the oligopolists each distort first downward and then upward on their interval of types served. Firm  $M$  distorts first further downward on a longer interval of low types and then further upward on a longer interval of high types. Hence, the surplus function offered by  $M$  is first flatter than in the oligopoly and then steeper, and so lies everywhere below it, since the two are by fiat equal at  $\theta_l$  and  $\theta_h$ . The proof takes into

<sup>37</sup>Note that firm  $M$ , in the exercise of its market power, may prefer to idle one or more of its technologies at the top or bottom. From the construction of  $\gamma^M$ , a nonempty subset of consecutive technologies will be active.

account that  $M$  will reallocate types across its constituent parts. In particular, the proof shows that, compared to oligopoly, constituent technologies in the monopoly below some threshold will serve higher intervals of types, while technologies in the monopoly above the threshold will serve lower intervals of types. In particular, the lowest and the highest constituent technologies will increase their market share.

### 5.3.3. *Oligopoly versus Monopoly With Endogenous Market Size*

Unless legally constrained to do so,  $M$  is unlikely to serve all of  $[\theta_l, \theta_h]$ . In oligopoly, then each firm, knowing that it would lose types at each end, was indifferent about decreasing the surplus by a small constant. But then the merged firm—which no longer suffers the loss of types at interior boundaries—strictly prefers to do so. By the next theorem, this remains true even after the merged firm has optimally reallocated actions and boundaries.

**THEOREM 6—Endogenous Span:** *Let  $M$  optimally serve  $[\theta_l^M, \theta_h^M]$ . If  $0 < \theta_l$ , then  $\theta_l < \theta_l^M$ , and if  $\theta_h < 1$ , then  $\theta_h^M < \theta_h$ . All types in  $(\theta_l, \theta_h)$  are strictly worse off compared to when  $M$  is forced to serve  $[\theta_l, \theta_h]$ , and so, a fortiori, are strictly worse off compared to oligopoly.*

Thus, the merged firm is both harder on the types it keeps and, except perhaps at the endpoints 0 and 1, strictly shrinks the set of types served.

### 5.3.4. *From an Oligopoly to an Oligopoly After a Merger*

Finally, let us see how these results help us to understand what happens when a subset of firms merges, creating an oligopoly with a smaller set of players. Consider first the case where a single firm controls a *nonsequential* set of technologies. Then, in equilibrium, the competing firms in the middle will be active, and we can consider IO and OB for each consecutive set of technologies separately.<sup>38</sup> So let firm  $M$  control a *sequential* set of technologies  $n_l, \dots, n_h$ .

First fix the behavior of firms outside of  $\{n_l, \dots, n_h\}$ . Then  $\bar{u}$ , which was assumed to be exogenous above, will now be determined by the best offer made by the firms controlling technologies below  $n_l$  and above  $n_h$ , and so by stacking will thus have the requisite shallow–steep property. Thus, by Theorem 6,  $M$  will wish to lower surplus to all types served and to cede market share.

It is intuitive that the full equilibrium should share these properties, with all firms offering less surplus than before and the merged firm losing share. But there are conflicting economic forces: when  $M$  maintains the set of types served, then as above,  $M$  moves the actions of its top and bottom agents closer to its nearest competitors, reducing its differentiation from them. Hence, when those competitors raise the surplus they are offering, they gain types faster, which pushes them to fight harder for types than before and *raises* the surplus they offer. On the other hand, as we argued, the merged firm has an incentive to shed market share to the boundary firms, who will then desire to *lower* surplus so as to profit from their larger span.

<sup>38</sup>To see this, recall first that our relevance assumption is equivalent to each  $V^n$  being somewhere strictly above the concave envelope of  $\max_{n' \neq n} V^{n'}$ , so it cannot happen that, by merging, nonconsecutive firms can turn a middle firm irrelevant by suitably mixing between their technologies, and recall that a merger is assumed to be “synergy-free,” in that if the merged firm holds fixed who it serves and how it serves them, then profits are the *same*.

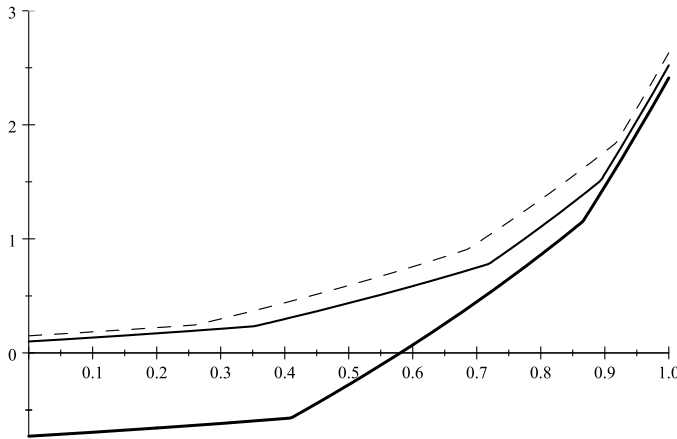


FIGURE 4.—A merger and a failing firm. The dashed locus is the equilibrium surplus from Figure 1. The medium weight locus is the equilibrium surplus when we put firms 2 and 3 under the control of a single firm,  $M$ , and the heavy locus is the surplus when firm 2 exits the market. Consumers, especially those of firm 1, are better off with the merger than with firm 2 failing.

One case where we can shed further light on this trade-off is when firms 2,  $\dots$ ,  $N$  merge into  $M$ , and thus the market becomes a duopoly with firm 1 and firm  $M$ . Consider the boundary type between the two firms. Pre-merger, firm 1 was facing firm 2 with  $\kappa^2 < 1$ , while now it is facing firm  $M$  with  $\kappa^M = 1$ . We show in Supplemental Material Section S5 that this increase in  $\kappa$  increases the market share of firm 1 unambiguously (the boundary type increases). Regarding the surplus offered by firm 1, the ambiguity described above remains. But, for the class of quadratic  $V^n$ 's, we show if firms are sufficiently differentiated, then firm 1 reduces the surplus of every type.

As a numeric illustration of the effects of merger, let us return to the four-firm setting of Figure 1. Figure 4 shows the effect first of merging firms 2 and 3, and then of instead eliminating firm 2.<sup>39</sup> Under the merger, all firms lower the surplus they offer to each type and the merged firm loses market share. Eliminating firm 2 is much worse for the agents, and so the *failing firm defense* is validated: it is better to let firm 2 be absorbed by firm 3 than to lose it altogether. The merged firm competes vigorously at both ends, while when firm 2 disappears, it serves to further differentiate firm 1 from its competitors. It is an interesting open question under what general conditions these intuitive comparative statics hold.

## 6. PROVING SUFFICIENCY AND EXISTENCE UNDER STACKING

We now outline the proof of Theorem 2. Recall that  $\Pi^n(\cdot, s^{-n})$  is not quasi-concave, since a convex combination of  $s^n$  and  $\hat{s}^n$  may win a set of types different from either of them. But then the first-order conditions need not imply optimality, complicating sufficiency. Existence is nontrivial because  $\Pi^n$  is not continuous at ties. And since  $\Pi^n(\cdot, s^{-n})$  is not quasi-concave, the set of best responses may be nonconvex, and the results of, for example, Reny (1999), need not apply.

In what follows next, we use stacking to move the analysis of  $n$ 's problem from choosing an action schedule and associated surplus function to two dimensions: each firm concen-

<sup>39</sup>For numerical analysis of the merger, one simply replaces the relevant instances of OB by (8).



trates simply on the choice of  $\theta_l^n$  and  $\theta_h^n$ , with the action profile tied down by IO. Later subsections analyze this problem and then use that analysis to prove sufficiency and existence. As a preview, the two dimensional characterization exhibits “enough” quasi-concavity to pin down the optimum via local conditions, thus ensuring sufficiency. The move also facilitates the proof of nonemptiness and convexity of the best-reply correspondence of each firm, which then helps us show existence via the Kakutani–Fan–Glicksberg theorem (Aliprantis and Border (2006, Corollary 17.55, p. 583)).

The first step of our attack is to restrict attention to menus that our necessary conditions suggest are reasonable:

**Cn1** The action function  $\alpha^n$  is continuous, with  $\alpha^n(\theta) \in [\gamma^n(\theta, 1), \gamma^n(\theta, 0)]$  for all  $\theta$  and  $n$ .<sup>40</sup>

**Cn2** The surplus function satisfies  $v^n \leq v^n_*$  for all  $n$ .

Fix  $n$  and  $s^{-n}$  satisfying conditions Cn1 and Cn2. By relevance, firm  $n$  earns strictly positive profits in any best response to  $s^{-n}$ . Let  $\theta^n_\times \in [0, 1]$ , where  $\times$  is mnemonic for “crossing,” be the point where  $v^{n-1}$  and  $v^{n+1}$  cross. By stacking, such a point exists with  $a^{-n} < \gamma^n(\cdot, 1)$  for  $\theta < \theta^n_\times$  and  $a^{-n} > \gamma^n(\cdot, 0)$  for  $\theta > \theta^n_\times$ . In Figure 1,  $\theta^n_\times$  is the point near 0.53 at which  $v^1$  and  $v^3$  cross.

One of the most important implications of stacking is that the difficult global property NP can be dropped from the analysis, since it is implied by the local conditions. Say that strategy  $s^n$  is *single-dominant* on  $(\tau_l, \tau_h)$  if  $v^n > v^{-n}$  on  $(\tau_l, \tau_h)$  and  $v^n < v^{-n}$  outside of  $[\tau_l, \tau_h]$ .

LEMMA 3—OB implies NP: *Assume stacking and let  $s$  be any strategy profile that satisfies Cn1 with the property that  $n$  wins with strictly positive probability. Then  $s^n$  is single-dominant on an interval with  $\theta^n_\times$  in its interior, and if  $s^n$  satisfies OB, then it satisfies NP.*

That  $s^n$  is single-dominant on an interval including  $\theta^n_\times$  follows since by Cn1 and stacking, the slope of  $v^n$  is strictly bigger than that of  $v^{-n}$  below  $\theta^n_\times$  and strictly smaller above it, and so  $v^n$  can only cross  $v^{-n}$  once below  $\theta^n_\times$  and once above, and these crossings are strict. And since  $n$  sometimes wins, it follows that  $v^n > v^{-n}$  on a nonempty interval that includes  $\theta^n_\times$ . That NP is redundant follows since by Cn1,  $a^{-n}(\theta)$  is above the efficient level for  $n$  for  $\theta > \theta_h$  and, hence, by (6), the profit to poaching is decreasing (recall that  $\pi_a \leq 0$  at  $\theta_h$ ). And, we show that near  $\theta_h$ , OB implies that  $n$  does not want to poach, a result which uses in an essential way that  $V^n$  is concave. The proof is similar for  $\theta < \theta_l$ .

Let us now move to two dimensions. Recall that  $\tilde{s}(\theta_l, \theta_h)$  solves the relaxed problem  $\mathcal{P}(\theta_l, \theta_h)$ , with action profile  $\gamma(\cdot, \tilde{\kappa}(\theta_l, \theta_h))$ , and  $\tilde{\kappa}(\theta_l, \theta_h) \in [H(\theta_l), H(\theta_h)]$ . Let  $r(\theta_l, \theta_h)$  be the value of  $\mathcal{P}(\theta_l, \theta_h)$ . We now relate the maximization of  $r$  to that of  $\Pi^n$ , the profit in the original problem. Recall that strategies are equivalent if they agree where they win.

PROPOSITION 1—Equivalence: *Assume stacking. Fix  $n$  and  $s^{-n}$  satisfying Cn1 and Cn2. Then  $r$  has a maximum  $(\theta_l, \theta_h)$ , and  $\hat{s}$  is a maximum of  $\Pi^n(\cdot, s^{-n})$  if and only if for some maximum  $(\theta_l, \theta_h)$  of  $r$ ,  $\hat{s}$  is single-dominant on  $(\theta_l, \theta_h)$ , and  $\hat{s}$  and  $\tilde{s}(\theta_l, \theta_h)$  are equivalent.*

Thus, each firm can simply choose the interval to serve, with the rest pinned down by IO. The proof proceeds as follows. We first show that  $r$  has a maximum at some point  $(\theta_l, \theta_h)$ , and the associated solution to the relaxed problem is feasible in the original game, serves the interval  $(\theta_l, \theta_h)$ , and has the same payoff as  $r$ . We then show that, for any strategy in

<sup>40</sup>We cannot just impose that firms use strategies of the form given by (1), as these do not form a convex subset of the strategy space.

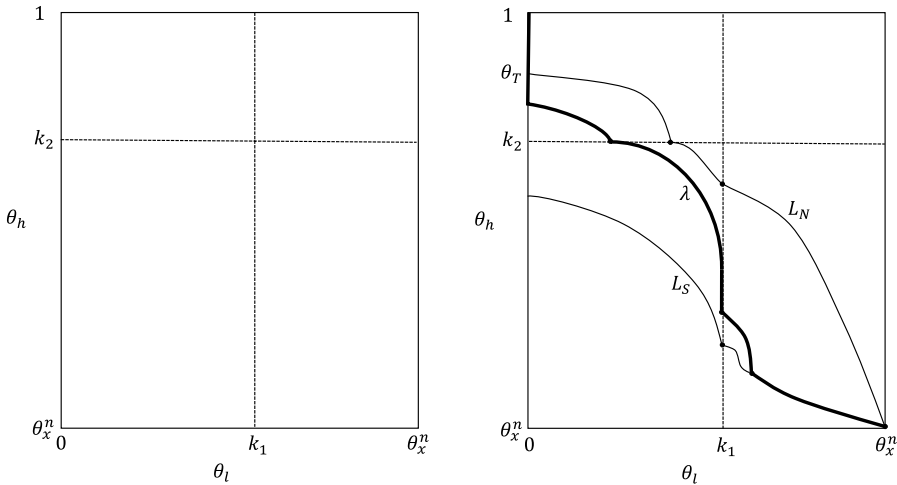


FIGURE 5.—The rectangle  $R$ . The left panel depicts the domain of  $r$ , kinks of  $v^{-n}$  at  $k_1, k_2$ , and  $\theta_x^n$ , and four rectangles where  $r$  is continuously differentiable. In the right panel, the area between  $L_S$  and  $L_N$  is  $\Theta$ . The thick line is the path described by  $\lambda$ . Where the path runs along  $L_S$ , we have  $r_{\theta_l} \leq 0$  and  $r_{\theta_h} > 0$ , and so  $\psi$  is increasing. The path never runs along  $L_N$ , where  $r_{\theta_l} < 0$ .

the original game, there is  $(\theta_l, \theta_h)$  such that  $r(\theta_l, \theta_h)$  is at least as big as the payoff to that strategy.

### 6.1. Unique Best Responses

In this section, we show that  $r$  has a *unique* maximum for any given  $s^{-n}$  satisfying Cn1 and Cn2, and that any critical point of  $r$  is that maximum. This is the crucial step in sufficiency and existence. We first develop some notation. Then we provide intuition and dive into the details.

We begin by showing that any optimum of  $r$  is in the rectangle  $R = [0, \theta_x^n] \times [\theta_x^n, 1]$  shown in Figure 5, left panel.<sup>41</sup> Let  $K$  be the set of points  $\theta$  where  $n$  transitions from one opponent to the next (note that  $|K| \leq N - 2$ ). In Figure 5, left panel,  $K = \{k_1, \theta_x^n, k_2\}$ , so that  $n$  has two competitors below it and two above. By Cn1, each point of discontinuity of  $\alpha^{-n}$  and, hence, each kink point of  $v^{-n}$ , is an element of  $K$ . Hence, letting  $\tilde{R}$  be any maximal rectangle on which  $n$ 's upper and lower opponents do not change (the four rectangles depicted in Figure 5, left panel),  $v^{-n}$  is continuously differentiable on  $\tilde{R}$  (it is convex with no kinks) and thus so is  $r$ .

#### 6.1.1. Hiking Toward a Proof

We need to show that  $r$  has a unique maximum characterized by first-order conditions corresponding to OB. We begin with some intuition. Fix the behavior of  $n$ 's opponents, and consider a landscape given by  $r$  on  $R$  (that is, the graph of the function  $r$  whose domain is  $R$ ), noting that in Figure 5,  $\theta_l$  is a choice from west to east, while  $\theta_h$  is a choice from south to north. This landscape has a complicated shape, with kinks and local minima, and places where  $r$  is negative, that is, “under water.” But when the other firms

<sup>41</sup>For firm 1,  $\theta_x^1 = 0$ , so the “rectangle”  $R$  becomes a vertical line segment and similarly, for firm  $N$ ,  $\theta_x^N = 1$ , so  $R$  becomes a horizontal line segment.

play strategies satisfying Cn1 and Cn2, the firm has available positive profit strategies. So consider the “islands” where  $r$  is positive.

We show first that on each island, any place where  $r$  is differentiable in one or both directions and where the appropriate first-order conditions are satisfied is also a local maximum in those directions. Thus, any local minimum where  $r$  is differentiable is in fact under water. Now fix any latitude (a choice of  $\theta_h$ ) with some land, and move from west to east. We show that despite the kinks in the landscape, there is a single interval of  $\theta_l$  above water, and payoffs are strictly quasi-concave on this interval. Hence, there is a unique highest point at each latitude.

Next we show that the set of latitudes where there is some land is an interval. But then there is a single island and a unique path running along the island where at each latitude, the corresponding point on the path is the highest point at that latitude. Finally, we show that payoffs are strictly quasi-concave as one hikes along this path. It follows that the island has a unique peak and that any point that satisfies the first-order conditions is in fact that maximum.<sup>42</sup>

### 6.1.2. Formalization and Outline of the Construction

Recall from Section 6 that  $\iota(\theta_l, \theta_h, \kappa) = v^{-n}(\theta_h) - v^{-n}(\theta_l) - \int_{\theta_l}^{\theta_h} \gamma(\tau, \kappa) d\tau$ , where since  $\gamma_\kappa < 0$ , we have  $\iota_\kappa > 0$ . Note also that on  $R$ ,  $\iota_{\theta_l} = -a^{-n}(\theta_l) + \gamma(\theta_l, \kappa) > 0$ , since  $\theta_l \leq \theta_x$ , and by Cn1 and stacking. Similarly,  $\iota_{\theta_h} > 0$  on  $R$ . Let the locus  $L_N$  be defined by  $\iota(\theta_l, \theta_h, H(\theta_l)) = 0$  and  $L_S$  be defined by  $\iota(\theta_l, \theta_h, H(\theta_h)) = 0$ . These are the north and south boundaries of  $\Theta = \{(\theta_l, \theta_h) \in R | \iota(\theta_l, \theta_h, \tilde{\kappa}(\theta_l, \theta_h)) = 0\}$ , and are depicted in Figure 5, right panel. The set  $\Theta$  will be central to our analysis, because we will see shortly that any maximum of  $r$  occurs either in  $\Theta$  or along a (specific) part of the boundary of  $R$ .

We begin by deriving the local properties of  $r$ . After some brush clearing, we show that on any given set  $\tilde{R}$  where  $n$ 's opponents do not change and for any given point in  $\tilde{R} \cap \Theta$ ,  $r_{\theta_l \theta_h} < 0$ . Further, if  $r_{\theta_l} = 0$ , then  $r$  is locally strictly concave in  $\theta_l$ . Similarly, if  $r_{\theta_h} = 0$ , then  $r$  is locally strictly concave in  $\theta_h$ , and anywhere that  $r_{\theta_l} = r_{\theta_h} = 0$ ,  $r$  is locally strictly concave in  $(\theta_l, \theta_h)$ . We use the local properties of  $r$  to analyze its maxima. We prove that on or below  $L_S$ , if  $r(\theta_l, \theta_h) > 0$ , then  $r_{\theta_h}(\theta_l, \theta_h) > 0$ , and on or above  $L_N$ , if  $r(\theta_l, \theta_h) > 0$ , then  $r_{\theta_l}(\theta_l, \theta_h) < 0$ . Assume first that, as in Figure 5, right panel,  $L_S$  hits the western boundary of  $R$ , let  $\theta_T \leq 1$  be the latitude at which  $L_N$  hits the boundary of  $R$ , and let  $A$  be the (possibly empty) segment of the western boundary of  $R$  above  $\theta_T$ . Then we show that any maximum of  $r$  occurs either in  $\Theta$ , with both the utility constraints (2) and (3) binding, or in  $A$ , with  $\theta_l = 0$  and (2) slack.

From here, we hike. For each  $\theta_h$ , let  $\Theta(\theta_h)$  be the interval of  $\theta_l$  such that  $(\theta_l, \theta_h) \in \Theta \cup A$ , so that for  $\theta'_h > \theta_T$ ,  $\Theta(\theta'_h) = \{0\}$ . Define  $\psi(\theta_h) = \max_{\theta_l \in \Theta(\theta_h)} r(\theta_l, \theta_h)$ , maximizing  $r$  moving east to west. Let  $D$  be the set of  $\theta_h$  such that  $\psi > 0$ . Fix  $\theta_h \in D$  with  $\theta_h < \theta_T$ . We show that  $r(\cdot, \theta_h)$  has a unique maximum  $\lambda(\theta_h)$ . Since  $r_{\theta_l \theta_h} < 0$  on each  $\tilde{R}$ ,  $\lambda$  is decreasing. The path described is  $(\lambda(\cdot), \cdot)$ , which runs northwest in Figure 5, right panel. We prove that  $\lambda$  is continuous, and hence so is  $\psi$ , and that  $D$  is an interval. The path  $\lambda$  never runs along  $L_N$ , because on  $L_N$ , profits strictly decrease in  $\theta_l$ . Where  $\lambda$  runs along  $L_S$ , we show that  $\psi$  strictly increases.

<sup>42</sup>Why did we not simply establish that  $r$  is strictly concave at any critical point? First,  $r$  can have local minima under water. Second, in  $\mathbb{R}^2$  this property of a function is not enough to establish uniqueness of its maximum. For example (Chamberland (2015, pp. 106–108)),  $f(x, y) = -(x^2 - 1)^2 - (x^2y - x - 1)^2$  has only two critical points: one at  $(-1, 0)$  and one at  $(1, 2)$ , both global maxima.

So consider any  $\hat{\theta}_h$  such that  $\lambda(\hat{\theta}_h)$  is in the interior of  $\Theta(\hat{\theta}_h)$ . We prove that the left and right derivatives of  $\psi$  at  $\hat{\theta}_h$  and the left and right partial derivatives of  $r$  with respect to  $\theta_h$  at  $(\lambda(\hat{\theta}_h), \hat{\theta}_h)$  agree. This follows from the envelope theorem. Using this, we show that  $\psi$  has a unique maximum on the interval  $D$ , that is, as one hikes northwest along  $\lambda$ . This uses the concavity properties of  $r$ , with the usual complexities at kink points. Finally, we show that if  $\theta_h^*$  is the unique maximizer of  $\psi$ , then  $(\lambda(\theta_h^*), \theta_h^*)$  is the unique maximizer of  $r$ .

Assume that instead of hitting  $R$ 's western boundary,  $L_S$  hits  $R$ 's northern boundary at  $(\tilde{\theta}_T, 1)$ . As before any optimum of  $r$  occurs either in  $\Theta$ , with both constraints binding, or on the segment of the northern boundary of  $R$  with  $\theta_l \leq \tilde{\theta}_T$ , with the constraint at 1 slack. We can then proceed as above, but exchange the roles of  $\theta_l$  and  $\theta_h$ , so that one defines  $\tilde{\lambda}(\theta_l)$  by first maximizing along north–south slices where  $\theta_l$  is held constant and then hikes along  $\tilde{\lambda}$ .

### 6.2. Sufficiency and Existence

Sufficiency in Theorem 2 follows intuitively since any strategy profile satisfying PS, IO, and OB corresponds for each  $n$  to a critical point of  $r^n$ , and so if Cn1 and Cn2 held, would be a best response by Proposition 1. We show how to modify strategies in an inessential way outside of the interval served so that Cn1 and Cn2 indeed hold. For existence, we further restrict the strategy space so that continuity holds and show that any equilibrium of the restricted game can be modified in an inessential way to be an equilibrium of the original game. The critical point in showing existence in the restricted game is that from above, any two best responses to a given strategy profile serve the same types and give them the same surplus. But then their convex combination does so too, and so is also a best response. We can then apply the Kakutani–Fan–Glicksberg theorem.

## 7. EXISTENCE BEYOND STACKING

Without stacking, existence in pure strategies becomes murkier. In this section, we prove existence in mixed strategies and discuss the challenges of pure-strategy existence without stacking. We will discard NEO and break all ties in favor of a firm that earns the highest profit.<sup>43</sup> For simplicity, we assume the agent has a constant outside option  $\bar{u} > -\infty$  and that the set of actions is a compact interval  $[0, \bar{a}]$ .<sup>44</sup> It is direct that Proposition 5 (Supplemental Material) holds with mixed strategies, so that cross-subsidization is not beneficial. So we restrict each firm to contracts that do not strictly lose money if accepted. This rules out the zero-profit equilibrium from footnote 17. We will show existence in the game where firms can randomize over such strategies.

To formalize, for any convex function  $j$ , let  $G(\theta, j) \subseteq [0, \bar{a}]$  be the subdifferential of  $j$  at  $\theta$ . Let  $q^n(\theta, v^n) = \max_{a \in G(\theta, v^n)} (V^n(a) + a\theta)$  be the most surplus that can be created for  $\theta$  without attracting another type.<sup>45</sup> Let  $\rho^n \equiv \min\{\bar{u} - 1, \min_{a \in [0, \bar{a}]} V^n(a)\} > \infty$ , noting that

<sup>43</sup>Similar to the discussion in footnote 15, it would be of interest to know whether the set of mixed-strategy equilibria of this game depends on the tie-breaking rule.

<sup>44</sup>This restriction would be justified in our original model if there is  $\bar{a} < \infty$  such that for all  $a > \bar{a}$ ,  $V^n(a) + a - \bar{u} < 0$ .

<sup>45</sup>The max operator is valid, since, as the Appendix establishes,  $G(\cdot, \cdot)$  is upper hemicontinuous.

there is no benefit to being able to offer surplus below  $\rho^n$ . Let

$$W^n = \left\{ v^n \left| \begin{array}{l} v^n \text{ is convex,} \\ v^n(\theta') - v^n(\theta) \in [0, \bar{a}(\theta' - \theta)] \text{ for all } \theta' \geq \theta \text{ in } [0, 1], \text{ and} \\ \rho^n \leq v^n(\theta) \leq q^n(\theta, v^n) \text{ for all } \theta \end{array} \right. \right\}$$

be the set of increasing and convex surplus functions for  $n$  with slope bounded by  $\bar{a}$ , and with surplus bounded below by  $\rho^n$  and above such that the firm does not lose money at any  $\theta$ .

Since  $v^n$  is convex,  $G(\cdot, v^n)$  is a singleton almost everywhere, so for any vector  $v \in W \equiv \prod_n W^n$ , there is no ambiguity in writing  $\Pi_e^n(v) = \int (q^n(\theta, v^n) - v^n(\theta)) \varphi_e^n(\theta, v) h(\theta) d\theta$ , where  $q^n - v^n$  plays the role of  $\pi^n$  and where  $\varphi_e^n$  is the efficient tie-breaking rule.<sup>46</sup> Let the mixed extension of  $(W^n, \Pi_e^n)_{n=1}^N$  be  $(\bar{W}^n, \bar{\Pi}_e^n)_{n=1}^N$ , where, for  $\mu \in \bar{W}$ ,  $\bar{\Pi}_e^n(\mu) = \int_W \Pi_e^n(v) d\mu(v)$  and we use the weak-\* topology on  $\bar{W}$ .

**THEOREM 7—Existence in Mixed Strategies:** *The set  $(\bar{W}^n, \bar{\Pi}_e^n)_{n=1}^N$  has an equilibrium.*

The proof (see Appendix D) uses Reny (1999, Corollary 5.2), where efficient tie-breaking is used to show reciprocal upper semicontinuity of payoffs in the mixed extension. The novel part of the proof, which may be of more general applicability, deals with the fact that a strategy “near”  $\mu^{-n}$  can with small probability be far from the support of  $\mu^{-n}$ .

Our intuition is that some suitable refinement (quite possibly different than NEO) will allow an existence result in pure strategies substantially broader than under stacking. Things get complicated, however, because if there are “support points”—binding offers by another firm in the middle of the interval where the firm is “always” winning—then the simple characterization provided by IO fails, and so the quasi-concavity that underlies our pure-strategy proof becomes much harder. That proof also relied on the strictly transversal nature of crossings.

One might also wonder about application of Reny (1999, Theorem 3.1) to establish pure-strategy existence. To do so would require quasi-concavity of payoffs in the strategy, which is complicated since payoffs for any given type  $\theta$  are quasi-concave, but not concave. Without special structure (Choi and Smith (2017), Quah and Strulovici (2012)), we do not see the path to showing that quasi-concavity is preserved under expectations, given that the set of types which the firm wins is changing as one mixes across strategies.

### 8. CONCLUSION

We extend the ubiquitous principal–agent problem in Mussa and Rosen (1978) and Maskin and Riley (1984) to a vertical oligopoly. Firms post menus to both screen agents and attract the right pool of types. We derive the equilibrium sorting, distortions, and gaps in quality/effort across firms. Under enough firm heterogeneity, a simple set of conditions is sufficient for a strategy profile to be an equilibrium, and an equilibrium exists. Contrary to monopoly, complete information can help the agents. We examine the model’s competitive limit and the effect of mergers.

Many extensions are worth pursuing. We conjecture that a more general interaction between the agent’s type and the match surplus generated will primarily present technical

<sup>46</sup>This reduction could have been made earlier, but it was convenient to make the action schedules explicit.

complications. It is important to extend sufficiency and pure-strategy existence when firms are less vertically differentiated, and to allow both horizontal and vertical differentiation. A pressing extension is to allow for common values and risk-averse agents, as in insurance markets.

APPENDIX A: PROOFS FOR SECTION 4.1

We show that each of the asserted properties in Theorem 1 holds (see also the Supplemental Material).

A.1. Proof for Section 4.1.1

PROPOSITION 2: *Each firm earns strictly positive profits in equilibrium.*

PROOF: By assumption for each  $n$ , there is an interval  $I$  such that  $v^n(\theta) > v_*^n(\theta)$  for all  $\theta \in I$ . Assume that on a positive-measure set of  $I$ ,  $v^{-n}(\theta) \geq v_*^n(\theta)$ . Then either some firm other than  $n$  is winning with positive probability and is losing money or  $n$  is winning, having offered surplus  $v^n(\theta) > v^{-n}(\theta) \geq v_*^n(\theta)$ , violating PP in either case.<sup>47</sup> Thus, for  $\varepsilon$  sufficiently small but positive, offering all types surplus  $v_*^n(\theta) - \varepsilon$  and action  $\alpha_*^n(\theta)$  earns at least  $\varepsilon$  on a positive-measure set of types and, hence,  $n$  must earn strictly positive profits in equilibrium. *Q.E.D.*

A.2. Proof for Section 4.1.2

PROPOSITION 3: *Let  $s$  be an equilibrium. Then, for all  $\theta$ ,  $v^o(\theta) \geq V^{(2)}(a^o(\theta)) + a^o(\theta)\theta$ .*

PROOF: Assume not. Then there exists  $\hat{\theta}$ , and two firms  $n'$  and  $n''$  such that for  $n \in \{n', n''\}$ ,  $V^n(a^o(\hat{\theta})) + a^o(\hat{\theta})\hat{\theta} - v^o(\hat{\theta}) > 0$ . Assume first that  $\hat{\theta} < 1$ . Then, since  $a^o$  is right-continuous, and  $v^o$  and  $V^n$  are continuous, there is  $\rho > 0$  such that for all  $\theta \in [\hat{\theta}, \hat{\theta} + \rho]$ ,  $V^n(a^o(\theta)) + a^o(\theta)\theta - v^o(\theta) > \rho$ . Let  $s^o = (a^o, v^o)$  and let  $P^{o,n} = \{\theta | \pi^n(\theta, s^o) \geq 0\}$ . Using Proposition 5, let  $\hat{s}^n = (\hat{\alpha}^n, \hat{v}^n)$  have  $\pi(\cdot, \hat{s}^n) \geq 0$  and agree with  $(a^o, v^o)$  on  $P^{o,n}$ . Let  $\hat{s}^n(\varepsilon) = (\hat{\alpha}^n, \hat{v}^n + \varepsilon)$ . Then, since  $\hat{s}^n$  and  $s^o$  agree on  $P^{o,n}$  and since  $v^o$  is the most anyone offers,  $\varphi^n(\theta, (\hat{s}^n(\varepsilon), s^{-n})) = 1$  on  $P^{o,n}$  for any  $\varepsilon > 0$ . Hence, since  $\pi^n(\cdot, \hat{\alpha}^n, \hat{v}^n) \geq 0$  and  $\pi_v^n = -1$ , we have

$$\Pi(\hat{s}^n(\varepsilon), s^{-n}) \geq -\varepsilon + \int_{P^{o,n}} \pi^n(\theta, s^o)h(\theta) d\theta.$$

Note next that on a full-measure set of  $\theta$  where  $\varphi^n > 0$ ,  $s^n = s^o$ . This follows since any time  $\varphi^n > 0$ ,  $v^n = v^o$ , and so if  $\varphi^n > 0$  on a positive-measure set, then  $\alpha^n = a^o$  almost everywhere on that set, since  $v^o$  is convex with derivative  $a^o$  almost everywhere. But then

$$\begin{aligned} \Pi(s^n, s^{-n}) &= \int \pi^n(\theta, s^n)\varphi^n(\theta, s)h(\theta) d\theta \\ &\leq \int_{P^{o,n}} \pi^n(\theta, s^n)\varphi^n(\theta, s)h(\theta) d\theta \\ &= \int_{P^{o,n}} \pi^n(\theta, s^o)\varphi^n(\theta, s)h(\theta) d\theta. \end{aligned}$$

<sup>47</sup>If  $n$  offers  $v^{-n}(\theta)$ , then firms other than  $n$  win with positive probability since ties are broken equiprobably.



Combining these two inequalities yields

$$\begin{aligned} \Pi(\hat{s}^n(\varepsilon), s^{-n}) - \Pi(s^n, s^{-n}) &\geq -\varepsilon + \int_{\rho O, n} \pi^n(\theta, a^O, v^O)(1 - \varphi^n(\theta, s))h(\theta) d\theta \\ &\geq -\varepsilon + \rho \int_{[\hat{\theta}, \hat{\theta} + \delta]} (1 - \varphi^n(\theta, s))h(\theta) d\theta, \end{aligned}$$

since on  $[\hat{\theta}, \hat{\theta} + \delta]$ ,  $\pi^n(\theta, a^O, v^O) \geq \rho > 0$ , and so, since  $\varepsilon$  is arbitrary,

$$\Pi(\hat{s}^n(\varepsilon), s^{-n}) - \Pi(s^n, s^{-n}) \geq \rho \int_{[\hat{\theta}, \hat{\theta} + \delta]} (1 - \varphi^n(\theta, s))h(\theta) d\theta.$$

But at any given  $\theta$ ,  $\varphi^{n'}(\theta, s) + \varphi^{n''}(\theta, s) \leq 1$ , and so the right-hand side (rhs) cannot be zero for both  $n'$  and  $n''$ . Hence, at least one of  $n'$  or  $n''$  has a strictly profitable deviation. The proof for  $\theta = 1$  is similar, simply working with a small neighborhood to the left of 1, and we are done. Q.E.D.

### A.3. Proof for Section 4.1.3

PROPOSITION 4: *Every Nash equilibrium (with or without NEO) has QPS.*

PROOF: Fix  $n$  and  $n' > n$ , let  $\theta_{\inf}^{n'}$  be the infimum of the support of  $\varphi^{n'}$ , and let  $\theta_{\sup}^n$  be the supremum of the support of  $\varphi^n$ . We will show that the only way that  $\theta_{\inf}^{n'} < \theta_{\sup}^n$  can hold is if  $n = n + 1$  and the two firms are tied at zero profits on  $(\theta_{\inf}^{n'}, \theta_{\sup}^n)$ . The core of the proof is to exploit that  $V^n$  is strictly supermodular in  $n$  and  $a$ .

Assume that  $\theta_{\inf}^{n'} < \theta_{\sup}^n$ . Conditional on  $\varphi^{n'}(\theta, s) > 0$ , with probability 1,  $\pi^{n'}(\theta, \alpha^{n'}, v^{n'}) \geq 0$  by PP and  $\pi^n(\theta, \alpha^{n'}, v^{n'}) \leq 0$  by NP. Hence, for any  $\varepsilon \in (0, (\theta_{\sup}^n - \theta_{\inf}^{n'})/2)$ , there is  $\theta_1 \in [\theta_{\inf}^{n'}, \theta_{\inf}^{n'} + \varepsilon]$ , where  $\varphi^{n'}(\theta_1) > 0$  and

$$\pi^{n'}(\theta_1, \alpha^{n'}, v^{n'}) \geq 0 \geq \pi^n(\theta_1, \alpha^{n'}, v^{n'}), \tag{9}$$

and, similarly, there is  $\theta_2 \in [\theta_{\sup}^n - \varepsilon, \theta_{\sup}^n]$ , where  $\varphi^n(\theta_2) > 0$  and

$$\pi^n(\theta_2, \alpha^n, v^n) \geq 0 \geq \pi^{n'}(\theta_2, \alpha^n, v^n). \tag{10}$$

By incentive compatibility, since  $\theta_2 > \theta_1$ , and since  $\varphi^{n'}(\theta_1) > 0$  and  $\varphi^n(\theta_2) > 0$ , it must be that  $\alpha^n(\theta_2) \geq \alpha^{n'}(\theta_1)$ . Adding (9) and (10), and cancelling common terms yields

$$V^{n'}(\alpha^{n'}(\theta_1)) + V^n(\alpha^n(\theta_2)) \geq V^n(\alpha^{n'}(\theta_1)) + V^{n'}(\alpha^n(\theta_2)).$$

Thus, since  $V^n(a)$  is strictly supermodular,  $\alpha^{n'}(\theta_1) = \alpha^n(\theta_2) \equiv \tilde{a}$ , and so, by incentive compatibility and since  $\varepsilon$  could be arbitrarily small,  $\alpha^{n'}(\theta) = \alpha^n(\theta) = \tilde{a}$  for all  $\theta \in (\theta_{\inf}^{n'}, \theta_{\sup}^n)$ . From (9),  $V^{n'}(\tilde{a}) \geq V^n(\tilde{a})$ , while from (10),  $V^{n'}(\tilde{a}) \leq V^n(\tilde{a})$ , and so  $V^{n'}(\tilde{a}) = V^n(\tilde{a}) \equiv \tilde{b}$ . But then, from (9),  $\pi^{n'}(\theta_1, \alpha^{n'}, v^{n'}) = 0$ , and from (10),  $\pi^n(\theta_2, \alpha^n, v^n) = 0$ . Finally, on  $(\theta_{\inf}^{n'}, \theta_{\sup}^n)$ ,  $(\pi(\theta, \alpha, v))_\theta = \pi_a(\theta, \alpha, v)\alpha_\theta(\theta) = 0$ , using  $v_\theta(\theta) = \alpha(\theta)$ . Hence,  $\pi^n = \pi^{n'} = 0$  on  $(\theta_{\inf}^{n'}, \theta_{\sup}^n)$ .

Now let us show that  $n' = n + 1$ . Assume that  $n' \neq n + 1$  and let  $n < n'' < n'$ . Assume first that  $V^{n''}(\tilde{a}) \leq \tilde{b} = V^n(\tilde{a})$ . Then since  $n'' > n$  and  $V^n(a)$  is strictly supermodular,  $V^{n''}(a) <$

$V^n(a)$  for all  $a < \tilde{a}$ , and similarly,  $V^{n'}(a) < V^n(a)$  for all  $a > \tilde{a}$ , contradicting that  $V^{n'}$  is somewhere uniquely maximal. Thus,  $V^{n'}(\tilde{a}) > \tilde{b}$  and so  $\pi^{n'}(\theta, \tilde{a}, v^{-n}) > 0$  on  $(\theta_{\inf}^{n'}, \theta_{\sup}^{n'})$ , which contradicts NP, since by definition of  $\theta_{\inf}^{n'}$  and  $\theta_{\sup}^{n'}$ ,  $\int_{\theta_{\inf}^{n'}}^{\theta_{\sup}^{n'}} (1 - \varphi^{n'})h > 0$ . Thus,  $n' = n + 1$  and  $\tilde{a} = a_e^n$ . Letting  $\theta_h^n = \theta_{\inf}^{n'}$  and  $\theta_l^{n+1} = \theta_{\sup}^{n'}$ , we have the claimed structure at ties. Finally, it must be that  $\theta_l^n < \theta_h^n$ , since by PP,  $n$  earns strictly positive expected profit, but on each type above  $\theta_h^n$  or below  $\theta_l^n$  either loses for sure or ties but earns 0. *Q.E.D.*

A.4. Proofs for Section 4.1.4

LEMMA 4: Fix  $n$ ,  $s^{-n}$ , and  $\hat{s}^n = (\hat{\alpha}, \hat{v})$ . If  $\hat{s}^n$  is a best response, then  $\hat{\alpha}$  must be continuous on any open interval where  $v^n \geq v^{-n}$ .

PROOF: Because  $\pi$  is strictly concave in  $a$ , any jump in  $\hat{\alpha}$  creates an opportunity for a strictly profitable perturbation. Let  $\hat{\alpha}$  jump from  $\underline{a}$  to  $\bar{a}$  at some point  $\theta_j$  belonging to an open interval where  $v^n \geq v^{-n}$ . Raise  $\hat{\alpha}$  by  $q$  on  $[\theta_j - \varepsilon, \theta_j)$  and lower it by  $q$  on  $[\theta_j, \theta_j + \varepsilon]$ , where for  $\varepsilon$  and  $q$  small enough, monotonicity is respected. This raises surplus slightly on  $(\theta_j - \varepsilon, \theta_j + \varepsilon)$  (by an amount at most  $q\varepsilon$ ), but otherwise does not affect  $v$ . The perturbed strategy serves with probability 1 on  $(\theta_j - \varepsilon, \theta_j + \varepsilon)$ , and any new type served by this perturbation makes at most a tiny loss, since by PP,  $\hat{s}^n$  loses money nowhere. We claim that because  $\pi$  is strictly concave in  $a$ , this perturbation is strictly profitable for sufficiently small  $\varepsilon$  and  $q$ , contradicting the optimality of  $\hat{s}^n$ . See the Supplemental Material for details. *Q.E.D.*

COROLLARY 1: Every Nash equilibrium that satisfies NEO has PS.

PROOF: Assume that for some  $n' > n$  and for some  $\hat{\theta} \in (\theta_l^n, \theta_h^n)$ ,  $v^{n'} = v^n$ . Then, since by NEO,  $\alpha^{n'} \geq a_e^{n'-1} \geq a_e^n \geq \alpha^n$  and, hence,  $v^{n'}(\theta) - v^n(\theta)$  is increasing,  $v^{n'} \geq v^n$  everywhere on  $[\hat{\theta}, \theta_h^n]$ , contradicting that  $n$  wins with probability 1 conditional on  $\theta \in (\theta_l^n, \theta_h^n)$ . *Q.E.D.*

APPENDIX B: PROOFS FOR SECTION 5

B.1. Proof for Section 5.1

PROOF OF THEOREM 3: Let us start from claim (ii). Let  $J^n$  be the set of types who are served by firm  $n$  under incomplete information and strictly prefer this to complete information. To see that  $J^n$  is a subset of  $[\theta_*^{n-1}, \theta_*^n]$ , consider any  $\theta \notin [\theta_*^{n-1}, \theta_*^n]$  that  $n$  serves. Then, since  $n$  earns strictly positive profits on all types served,  $v^n(\theta) < v_*^n(\theta) \leq v_*^{(2)}(\theta)$ , the second-order statistic on  $\{v_*^{n'}\}_{n' \in \{1, \dots, N\}}$ . But  $\theta$  receives  $v_*^{(2)}(\theta)$  under complete information and so  $\theta \notin J^n$ .

Next note that on  $[\theta_*^{n-1}, \theta_*^n]$ ,  $v_*^{(2)}(\theta) = \max_{n' \neq n} v_*^{n'}(\theta)$ . By stacking,  $v_*^{(2)}$  on  $[\theta_*^{n-1}, \theta_*^n]$  is first shallow and then steep relative to  $\gamma^n(\cdot, \kappa)$  for any  $\kappa \in [0, 1]$ . Hence, if  $J^n$  is nonempty, then  $v^n$  must cross  $v_*^{(2)}$  exactly twice, once where  $v_*^{(2)}$  is shallow and once where it is steep. Thus,  $J^n$  is an interval, where if  $n \notin \{1, N\}$ , then  $J^n$  is of the form  $(\underline{J}^n, \bar{J}^n)$ , if  $J^1$  is nonempty, it is of the form  $[0, \bar{J}^1]$ , and if  $J^N$  is nonempty, it is of the form  $(\underline{J}^N, 1]$ .

Let us turn to claim (i). On  $(\bar{J}^n, \underline{J}^{n+1})$ , the agent by definition strictly prefers complete information. Since by claim (ii),  $n$  was the uniquely most efficient firm everywhere on  $J^n$  and  $n + 1$  was the uniquely most efficient firm everywhere on  $J^{n+1}$ , we have  $\theta_*^n \in [\bar{J}^n, \underline{J}^{n+1}]$ .

But  $v_*^{(2)}(\theta_*^n) = v_*^{(1)}(\theta_*^n) > v^O(\theta_*^n)$ , where  $v^O(\theta)$  is the surplus of type  $\theta$  with incomplete information, since each firm makes strictly positive profits on all types in equilibrium. Hence,  $\theta_*^n \in (\bar{J}^n, \underline{J}^{n+1})$ , which is thus nonempty. Let  $\theta^n$  be the boundary between firms  $n$  and  $n + 1$  under incomplete information. Then  $\theta^n \in [\bar{J}^n, \underline{J}^{n+1}]$ , since by construction  $n$  serves types in  $J^n$  and  $n + 1$  serves types in  $J^{n+1}$ . But since  $v^n(\theta^n) = v^{n+1}(\theta^n)$  and since each firm is making strict profits,  $v^n(\theta^n) < v_*^{(2)}(\theta^n)$ , and, thus,  $\theta^n \in (\bar{J}^n, \underline{J}^{n+1})$ . Finally, note that at  $\theta_*^n$ , both firms earn strictly positive profits under incomplete information, but zero under complete information.

We prove (iii) in the Supplemental Material by constructing an example where all types prefer complete information. Q.E.D.

### B.2. Proofs for Section 5.3

Let the span of the constituent firms in  $M$  pre-merger be  $[\theta_l, \theta_h]$ . Let  $v^O$  be the surplus being offered on  $[\theta_l, \theta_h]$  pre-merger. Our first key step is to understand how the merged firm adjusts which types are served by which firm. Let  $\theta^{O,n}$  ( $O$  for oligopoly) be the boundary between firms  $n$  and  $n + 1$  in the pre-merger equilibrium, where  $\theta^{O,n_l-1} = \theta_l$  and  $\theta^{O,n_h} = \theta_h$ . Motivated by (8), consider the function  $q$  given by

$$q(k^n, k^{n+1}, \theta) = V^n(\gamma^n(\theta, k^n)) + \theta\gamma^n(\theta, k^n) - (V^{n+1}(\gamma^{n+1}(\theta, k^{n+1})) + \theta\gamma^{n+1}(\theta, k^{n+1})) + \frac{k^n - H(\theta)}{h(\theta)}(\gamma^{n+1}(\theta, k^{n+1}) - \gamma^n(\theta, k^n)).$$

Note that by the definition of  $\gamma^n$ ,  $V_a^n(\gamma^n(\theta, k^n)) + \theta = (k^n - H(\theta))/h(\theta)$ , and so  $q_\theta(k, k, \theta) = ((k - H(\theta))/h(\theta))_\theta - 1)(\gamma^{n+1}(\theta, k) - \gamma^n(\theta, k)) < 0$ . Thus,  $\theta^{M,n}$  is the unique solution to  $q(\kappa^M, \kappa^M, \theta^{M,n}) = 0$ , where  $\theta^{M,n} = 0$  if  $q(\kappa^M, \kappa^M, 0) < 0$ , and  $\theta^{M,n} = 1$  if  $q(\kappa^M, \kappa^M, 1) > 0$ , and where, as noted before,  $\theta^{M,n}$  depends on  $\kappa$  but not on  $\bar{u}$ . Thus, given  $\kappa^M$  and given that the merged firm chooses to serve type  $\theta$ , it does so optimally with firm  $n_l$  for  $\theta$  below  $\theta^{M,n_l}$ , firm  $n \in \{n_l, \dots, n_h - 1\}$  for  $\theta$  between  $\theta^{M,n-1}$  and  $\theta^{M,n}$ , and firm  $n_h$  for  $\theta$  above  $\theta^{M,n_h-1}$ .

LEMMA 5: *If  $\kappa^M \geq \kappa^{n+1}$ , then  $\theta^{M,n} > \theta^{O,n}$ , while if  $\kappa^M \leq \kappa^n$ , then  $\theta^{M,n} < \theta^{O,n}$ .*

PROOF: Fix  $n$ , let  $\theta^O = \theta^{O,n}$ , and let  $\theta^M = \theta^{M,n}$ . Firm  $n$ 's optimal boundary condition in the oligopoly is

$$V^n(\gamma^n(\theta^O, \kappa^n)) + \theta^O\gamma^n(\theta^O, \kappa^n) - v^O(\theta^O) + \frac{\kappa^n - H(\theta^O)}{h(\theta^O)}(\gamma^{n+1}(\theta^O, \kappa^{n+1}) - \gamma^n(\theta^O, \kappa^n)) = 0,$$

while  $V^{n+1}(\gamma^{n+1}(\theta^O, \kappa^{n+1})) + \theta^O\gamma^{n+1}(\theta^O, \kappa^{n+1}) - v^O(\theta^O) > 0$  (by the discussion following Lemma 2). Subtracting and cancelling  $v^O$ ,  $q(\kappa^n, \kappa^{n+1}, \theta^O) < 0$ . But for  $k^{n+1} > k^n$ , we have  $q_{k^n}(k^n, k^{n+1}, \theta) = (1/h(\theta))(\gamma^{n+1}(\theta, k^{n+1}) - \gamma^n(\theta, k^n)) > 0$  and  $q_{k^{n+1}}(k^n, k^{n+1}, \theta) = ((k^n - k^{n+1})/h(\theta))\gamma_k^{n+1}(\theta, k^{n+1}) > 0$ , using  $k^{n+1} > k^n$ . Thus, if  $\kappa^M \leq \kappa^n$ , then  $0 > q(\kappa^n, \kappa^{n+1}, \theta^O) \geq q(\kappa^M, \kappa^{n+1}, \theta^O) > q(\kappa^M, \kappa^M, \theta^O)$ , and so, since  $q_\theta(k, k, \theta) < 0$  and since  $q(\kappa^M, \kappa^M, \theta^{M,n}) = 0$ , it follows that  $\theta^{M,n} < \theta^{O,n}$ .

Let us turn to  $\kappa^M \geq \kappa^{n+1}$ . Firm  $n + 1$ 's optimal boundary condition for  $\theta^O$  can be written

$$\begin{aligned}
 & - (V^{n+1}(\gamma^{n+1}(\theta^O, \kappa^{n+1})) + \theta^O \gamma^{n+1}(\theta^O, \kappa^{n+1}) - v^O(\theta^O)) \\
 & + \frac{\kappa^{n+1} - H(\theta^O)}{h(\theta^O)} (\gamma^{n+1}(\theta^O, \kappa^{n+1}) - \gamma^n(\theta^O, \kappa^n)) = 0,
 \end{aligned}$$

and so adding  $V^n(\gamma^n(\theta^O, \kappa^n)) + \theta^O \gamma^n(\theta^O, \kappa^n) - v^O(\theta^O) > 0$  to the left-hand side (lhs), and adding and subtracting  $\kappa^n$  in the term  $\kappa^{n+1} - H(\theta^O)$ , we arrive at  $\hat{q}(\kappa^n, \kappa^{n+1}, \theta^O) > 0$ , where

$$\hat{q}(\kappa^n, \kappa^{n+1}, \theta) \equiv q(\kappa^n, \kappa^{n+1}, \theta) + \frac{\kappa^{n+1} - \kappa^n}{h(\theta)} (\gamma^{n+1}(\theta, \kappa^{n+1}) - \gamma^n(\theta, \kappa^n)).$$

Using the expressions for the derivatives of  $q$  from above, we have that for  $\kappa^{n+1} > \kappa^n$ ,  $\hat{q}_{k^n}(\kappa^n, \kappa^{n+1}, \theta) = -\frac{\kappa^{n+1} - \kappa^n}{h(\theta)} (\gamma'_\kappa(\theta, \kappa^n)) > 0$ ,  $\hat{q}_{k^{n+1}}(\kappa^n, \kappa^{n+1}, \theta) = \frac{1}{h(\theta)} (\gamma^{n+1}(\theta, \kappa^{n+1}) - \gamma^n(\theta, \kappa^n)) > 0$  and  $\hat{q}_\theta(\kappa, \kappa, \theta) = q_\theta(\kappa, \kappa, \theta) < 0$ . Similarly,  $\hat{q}_\theta(k, k, \theta) = q_\theta(k, k, \theta) < 0$ . We thus have  $0 < \hat{q}(\kappa^n, \kappa^{n+1}, \theta^O) \leq \hat{q}(\kappa^n, \kappa^M, \theta^O) < \hat{q}(\kappa^M, \kappa^M, \theta^O)$  and so  $\theta^M > \theta^O$ . Q.E.D.

Our next lemma shows that the monopolist always uses a surplus function that is “more convex” than the oligopolists’. Let  $\hat{v}^M$  be the optimal surplus function when the merged firm  $M$  is forced to serve exactly  $[\theta_l, \theta_h]$ . Say that a continuous function  $g$  on  $[0, 1]$  is a *tent* on  $[\underline{\theta}, \bar{\theta}] \subseteq [0, 1]$  if there are  $\underline{\theta} \leq \tau_l \leq \tau_h \leq \bar{\theta}$  such that  $g$  is strictly increasing on  $[\underline{\theta}, \tau_l)$ , constant on  $[\tau_l, \tau_h]$ , and strictly decreasing on  $(\tau_h, \bar{\theta}]$ , where at least one of  $\tau_l > \underline{\theta}$  or  $\tau_h < \bar{\theta}$  holds.

LEMMA 6: *The functions  $v^O - v^M$  and  $v^O - \hat{v}^M$  are tents on  $[\theta_l, \theta_h]$ . If  $\kappa^M \geq \kappa^{n_h}$ , then  $v^O - v^M$  is increasing, and if  $\kappa^M \leq \kappa^{n_l}$ , then  $v^O - v^M$  is decreasing.*

PROOF: We provide the proof for  $v^M$ . The proof for  $\hat{v}^M$  is the same, choosing boundaries between firms in the monopoly to reflect the  $\kappa$  chosen by  $M$  when it is forced to maintain its span.

Consider the case  $\kappa^{n_l} < \kappa^M < \kappa^{n_h}$ , so that there is  $n^* \in \{n_l, \dots, n_h - 1\}$  (not necessarily unique) with  $\kappa^{n^*} \leq \kappa^M \leq \kappa^{n^*+1}$ . Assume first that  $\theta^{M, n^*} \geq \theta^{O, n^*}$ . Then, by the contrapositive to the relevant part of Lemma 5, we have  $\kappa^M > \kappa^{n^*}$ . Furthermore, for any  $n \leq n^* - 1$  (if there are any), we have  $\theta^{M, n} > \theta^{O, n^*}$ , again by Lemma 5. Thus, for any  $\theta \in [\theta_l, \theta^{M, n^*})$ , we have that the active firm in the monopoly at  $\theta$  has a weakly lower index than in the oligopoly, and that for whatever firm is operating, since  $\kappa^M > \kappa^n$ , the monopolist firm of that index takes a strictly lower action than the oligopolist of the same index. Hence,  $v_\theta^O > v_\theta^M$  on the interval  $[\theta_l, \theta^{M, n^*})$ . Note also that since  $\theta^{M, n^*} \geq \theta^{O, n^*} > \theta_l$ ,  $[\theta_l, \theta^{M, n^*})$  is nonempty.

Note next that for  $n \geq n^* + 1$ , since  $\kappa^M \leq \kappa^{n^*+1} \leq \kappa^n$ , by Lemma 5,  $\theta^{M, n} < \theta^{O, n}$ . On  $(\theta^{M, n^*}, \theta^{M, n^*+1})$  both the oligopoly and the monopolist are using firm  $n^* + 1$ , but  $v_\theta^O \leq v_\theta^M$  since  $\kappa^M \leq \kappa^{n^*+1}$ . Finally on  $(\theta^{M, n^*+1}, \theta_h]$ , the active firm in monopoly has a weakly greater index than in oligopoly and is acting according to a strictly lower  $\kappa$ , since the relevant  $\kappa$  in oligopoly is at least  $\kappa^{n^*+2} > \kappa^{n^*+1} \geq \kappa^M$ . Hence,  $v_\theta^O < v_\theta^M$  on  $(\theta^{M, n^*+1}, \theta_h]$ .

Assume next that  $\theta^{M, n^*} < \theta^{O, n^*}$ . As above, for  $n \leq n^* - 1$ , by Lemma 5,  $\theta^{M, n} > \theta^{O, n}$ , since  $\kappa^M \geq \kappa^{n^*} = \kappa^{(n^*-1)+1}$ . Thus, on  $[\theta_l, \theta^{M, n^*-1})$ , the active firm in monopoly has a weakly lower

index than in oligopoly, and acts according to  $\kappa^M \geq \kappa^{n^*} > \kappa^n$  for whatever firm is active in monopoly. Hence,  $v_\theta^O > v_\theta^M$ . On the interval  $\theta \in (\theta^{M,n^*-1}, \theta^{M,n^*})$  both the oligopoly and the monopoly are using firm  $n^*$ , and so, since  $\kappa^M \geq \kappa^{n^*}$ ,  $v_\theta^O \geq v_\theta^M$ . Finally, by Lemma 5  $\kappa^M < \kappa^{n^*+1}$ , and so, for  $n \geq n^* + 1$ ,  $\theta^{M,n} < \theta^{O,n}$ . Hence, for all  $\theta \in (\theta^{M,n^*}, \theta_h]$ , and as above,  $v_\theta^O < v_\theta^M$ , where again, since  $\theta^{M,n^*} \leq \theta^{O,n^*} < \theta_h$ ,  $(\theta^{M,n^*}, \theta_h]$  is nonempty.

Consider next  $\kappa^M \leq \kappa^{n_l}$ . By Lemma 5,  $\theta^{M,n} < \theta^{O,n}$  for all  $n_l \leq n \leq n_h - 1$ . Thus, as above,  $v_\theta^O \leq v_\theta^M$  on  $[\theta_l, \theta_h]$  and strictly so on the nonempty interval  $[\theta^{M,n_l}, \theta_h]$ . Similarly, if  $\kappa^M \geq \kappa^{n_h}$ , then, by Lemma 5,  $\theta^{M,n} > \theta^{O,n}$  for all  $n \in \{n_l, \dots, n_h - 1\}$ , and so  $v_\theta^O > v_\theta^M$  everywhere on  $[\theta_l, \theta_h]$  and strictly so on the nonempty interval  $[\theta_l, \theta^{M,n_h-1}]$ . Q.E.D.

PROOF OF THEOREM 5: By Lemma 6,  $v^O - \hat{v}^M$  is a tent, where, since  $v^O(\theta_l) - \hat{v}^M(\theta_l) = v^O(\theta_h) - \hat{v}^M(\theta_h) = 0$ , both the strictly increasing and strictly decreasing intervals of  $v^O - \hat{v}^M$  are nonempty and, hence,  $\hat{\kappa}^M \in (\kappa^{n_l}, \kappa^{n_h})$ . Q.E.D.

PROOF OF THEOREM 6: Let  $n'_l \geq n_l$  and  $n'_h \leq n_h$  be the lowest and highest active firms in monopoly. Assume that  $\theta_h$  is interior and that  $\theta_h^M \geq \theta_h$ . We will show a contradiction. Assume first that  $\kappa^M \geq \kappa^{n'_h}$ . Then, for all  $n \in \{n'_l, \dots, n'_h - 1\}$ ,  $\theta^{M,n} > \theta^{O,n}$ . But then, by definition of  $n'_h$  and using stacking,  $v_\theta^O(\theta) > v_\theta^M(\theta)$  everywhere. Hence, since  $\theta_h^M \geq \theta_h$ , which implies  $v^M(\theta_h) \geq v^O(\theta_h)$ , we have  $v^M(\theta_l) > v^O(\theta_l)$  and, hence,  $\theta_l^M < \theta_l$ . Let  $\theta^*$  be the lower boundary of firm  $n'_l$  in oligopoly. If  $n'_l = n_l$ , let  $\tilde{u}(\theta) = \bar{u}(\theta)$  for all  $\theta$ . If  $n'_l > n_l$ , then let  $\tilde{u} = \bar{u}$  for  $\theta < \theta_l$  and define  $\tilde{u} = \max_{n \in \{n_l, \dots, n'_l-1\}} v^n$  for  $\theta > \theta_l$ . As a maximum of convex functions,  $\tilde{u}$  is convex, and by stacking,  $\tilde{u}' < \gamma^{n'_l}$  for any  $\theta$  and  $\kappa \in [0, 1]$ . Thus, from firm  $n'_l$ 's optimal boundary condition at the bottom

$$0 = \omega^{n'_l}(\theta^*, \kappa^{n'_l}) \equiv V^{n'_l}(\gamma^{n'_l}(\theta^*, \kappa^{n'_l})) + \theta^* \gamma^{n'_l}(\theta^*, \kappa^{n'_l}) - \tilde{u}(\theta^*) + \frac{\kappa^{n'_l} - H(\theta^*)}{h(\theta^*)} (\tilde{u}'(\theta^*) - \gamma^{n'_l}(\theta^*, \kappa^{n'_l})).$$

Note that  $\omega_\theta^{n'_l}(\theta, \kappa^{n'_l}) = ((\kappa^{n'_l} - H/h)_\theta - 1)(\tilde{u}' - \gamma^{n'_l}) + ((\kappa^{n'_l} - H)/h)\tilde{u}'' > 0$  for all  $\theta \leq \theta^*$ , using that  $((\kappa^{n'_l} - H)/h)_\theta - 1 < 0$ ,  $\tilde{u}' - \gamma^{n'_l} < 0$ , and  $\kappa^{n'_l} - H(\theta) \geq \kappa^{n'_l} - H(\theta^*) > 0$  by OB, and  $\tilde{u}'' \geq 0$ . Hence,  $\omega^{n'_l}(\theta_l^M, \kappa^{n'_l}) < 0$ . But,  $\omega_\kappa^{n'_l} = (\tilde{u}' - \gamma^{n'_l})/h < 0$  and, hence,  $\omega^{n'_l}(\theta_l^M, \kappa^M) < 0$  since  $\kappa^M \geq \kappa^{n'_h} \geq \kappa^{n'_l}$ . Since  $\tilde{u}$  and  $\bar{u}$  coincide at  $\theta_l^M < \theta_l$ , this contradicts that  $\theta_l^M$  is optimal.

Hence, we must have  $\kappa^M < \kappa^{n'_h}$ . Now let  $\theta^*$  be the upper boundary of firm  $n'_h$  in oligopoly. If  $n'_h = n_h$ , let  $\tilde{u} = \bar{u}$ , while if  $n_h > n'_h$ , then let  $\tilde{u} = \bar{u}$  for  $\theta > \theta_h$  and define  $\tilde{u} = \max_{n \in \{n'_h+1, \dots, n_h\}} v^n$  for  $\theta < \theta_h$ . Firm  $n'_h$ 's optimal boundary condition at  $\theta^*$  in the oligopoly is

$$0 = \omega^{n'_h}(\theta^*, \kappa^{n'_h}) \equiv V^{n'_h}(\gamma^{n'_h}(\theta^*, \kappa^{n'_h})) + \theta^* \gamma^{n'_h}(\theta^*, \kappa^{n'_h}) - \tilde{u}(\theta^*) + \frac{\kappa^{n'_h} - H(\theta^*)}{h(\theta^*)} (\tilde{u}'(\theta^*) - \gamma^{n'_h}(\theta^*, \kappa^{n'_h})).$$

Note that  $\theta^* \leq \theta_h \leq \theta_h^M$ . But then, for all  $\theta \in [\theta^*, \theta_h^M]$ ,

$$\omega_\theta^{n'_h}(\theta, \kappa^{n'_h}) = \left( \left( \frac{\kappa^{n'_h} - H}{h} \right)_\theta - 1 \right) (\tilde{u}' - \gamma^{n'_h}) + \frac{\kappa^{n'_h} - H}{h} \tilde{u}'' < 0,$$

since  $((\kappa^{n'_h} - H)/h)_\theta - 1 < 0$ ,  $\bar{u}' - \gamma^{n'_h} > 0$ , and  $\kappa^{n'_h} - H(\theta) \leq \kappa^{n'_h} - H(\theta^*) < 0$  by OB, and  $\tilde{u}' \geq 0$ . Hence, since  $0 = \omega^{n'_h}(\theta^*, \kappa^{n'_h})$ , we have  $0 \geq \omega^{n'_h}(\theta_h^M, \kappa^{n'_h})$ . Finally, note that  $\omega_{\kappa}^{n'_h} = (\tilde{u}' - \gamma^{n'_h})/h > 0$ , since we are on the steep part of  $\bar{u}$ . Thus, since  $\kappa^M < \kappa^{n'_h}$ , we have  $0 > \omega^{n'_h}(\theta_h^M, \kappa^M)$ . This contradicts that the merged firm has set  $\theta_h^M$  optimally. We thus have a contradiction to  $\theta_h^M \geq \theta_h$ , proving that  $\theta_h^M < \theta_h$ . Similarly,  $\theta_l^M > \theta_l$ , so that the merged firm strictly sheds market share at each end.

Let  $[\theta_l^M, \theta_h^M]$  be the span of the merged firm, with associated  $v^M$  and  $\kappa^M$ . Let  $\hat{\kappa}^M$  govern the action when the span is constrained to be  $[\theta_l, \theta_h]$ . We will show that  $\delta(\theta) = \hat{v}^M(\theta) - v^M(\theta)$  is everywhere strictly positive. Given Theorem 5, we then have  $v^O > \hat{v}^M > v^M$ , establishing the result. Note first that since  $\bar{u}$  is shallow-steep,

$$v^M(\theta_l^M) = \bar{u}(\theta_l^M) = \bar{u}(\theta_l) + \int_{\theta_l}^{\theta_l^M} \bar{u}'(\tau) d\tau < \bar{u}(\theta_l) + \int_{\theta_l}^{\theta_l^M} \gamma^M(\tau, \hat{\kappa}^M) d\tau = \hat{v}^M(\theta_l^M),$$

and so  $\delta(\theta_l^M) > 0$ . Similarly (integrating from  $\theta_h^M$  to  $\theta_h$ ),  $\delta(\theta_h^M) > 0$ . But, for all  $\theta \in (\theta_l^M, \theta_h^M)$ ,  $\delta_\theta(\theta) = \gamma^M(\theta, \kappa^M) - \gamma^M(\theta, \hat{\kappa}^M)$ , and so if  $\kappa^M \geq \hat{\kappa}^M$  ( $\kappa^M \leq \hat{\kappa}^M$ ), then  $\delta$  is monotone decreasing (increasing) on  $[\theta_l^M, \theta_h^M]$ . But then, because  $\delta > 0$  at each end point,  $\delta > 0$  is strictly positive everywhere on  $[\theta_l^M, \theta_h^M]$  and we are done. Q.E.D.

### APPENDIX C: PROOFS FOR SECTION 6

The main text lays out the development. Results stated in the text are proved in sequential order. *For this and the next two subsections, we assume stacking, and whenever we fix  $n$  and  $s^{-n}$ , we assume  $s^{-n}$  satisfies Cn1 and Cn2. We omit the superscript  $n$  wherever possible.*

**PROOF OF LEMMA 3:** Given the discussion in the main text, it remains to show that poaching just above  $\theta_h$  does not make sense. But from (5), since  $\pi$  is strictly concave in  $a$ , we have that  $0 = \pi_a(\theta_h, \alpha(\theta_h), v)(a^{-n}(\theta_h) - \alpha(\theta_h)) + \pi(\theta_h, \alpha, v) > \pi(\theta_h, a^{-n}(\theta_h), v) - \pi(\theta_h, \alpha, v) + \pi(\theta_h, \alpha, v) = \pi(\theta_h, a^{-n}, v) = \pi(\theta_h, a^{-n}, v^{-n})$ , where the last equality follows from  $v(\theta_h) = v^{-n}(\theta_h)$ . Q.E.D.

#### C.1. Proving Proposition 1

**COROLLARY 2:** *Assume that  $\theta_l < \theta_h$  and  $r(\theta_l, \theta_h) \geq 0$ , and let  $\tilde{s}(\theta_l, \theta_h) = (\tilde{\alpha}, \tilde{v})$ . If  $\tilde{\kappa}(\theta_l, \theta_h) = H(\theta_h)$ , then  $\pi(\theta_h, \tilde{\alpha}, \tilde{v}) > 0$ , and if  $\tilde{\kappa}(\theta_l, \theta_h) = H(\theta_l)$ , then  $\pi(\theta_l, \tilde{\alpha}, \tilde{v}) > 0$ .*

**PROOF:** If  $\tilde{\kappa}(\theta_l, \theta_h) = H(\theta_h)$ , then by Lemma 2,  $\pi(\cdot, \tilde{\alpha}, \tilde{v})$  is strictly single-peaked with peak at  $\theta_h$ . Hence, if  $\pi(\theta_h, \tilde{\alpha}, \tilde{v}) \leq 0$ , then  $\pi(\theta, \tilde{\alpha}, \tilde{v}) < 0$  for all  $\theta < \theta_h$  and so  $r(\theta_l, \theta_h) < 0$ , and similarly for  $\tilde{\kappa}(\theta_l, \theta_h) = H(\theta_l)$ . Q.E.D.

**LEMMA 7:** *The function  $r$  has a maximum, and at any maximum  $(\theta_l, \theta_h)$ , (i)  $\tilde{s}(\theta_l, \theta_h) \in S$ , (ii) if  $\theta_l > 0$ , then  $\tilde{v}(\theta_l) = v^{-n}(\theta_l)$  and if  $\theta_h < 1$ , then  $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$ , and (iii)  $\tilde{s}(\theta_l, \theta_h)$  is single-dominant on  $(\theta_l, \theta_h)$  with  $\Pi(\tilde{s}(\theta_l, \theta_h), s^{-n}) = r(\theta_l, \theta_h)$ .*

**PROOF:** Note that  $r$  is continuous, since  $\tilde{\kappa}$  is continuous in  $(\theta_l, \theta_h)$ ,  $\gamma$  is continuous in  $\kappa$ ,  $\tilde{v}$  is continuous in  $(\theta_l, \theta_h, \kappa)$ , and the integral in the objective function is continuous in its endpoints. Since  $\{\theta_l, \theta_h | 0 \leq \theta_l \leq \theta_h \leq 1\}$  is compact,  $r$  has a maximum. Part (i) follows



since  $\tilde{\kappa} \in [0, 1]$ , and so, from footnote 22,  $\tilde{\alpha}$  is monotone, and since  $\tilde{v}(\theta) = \tilde{v}(0) + \int_0^\theta \tilde{\alpha} d\tau$  by construction.

To see (ii), consider any maximizer  $(\theta_l, \theta_h)$  of  $r$  at which  $\tilde{v}(\theta_h) > v^{-n}(\theta_h)$ . We will show that  $r_{\theta_h}(\theta_l, \theta_h) > 0$ , which, since  $(\theta_l, \theta_h)$  is optimal, implies  $\theta_h = 1$ . To do so, note first that for all  $\theta'_h$  on a neighborhood of  $\theta_h$ ,  $\tilde{\kappa}(\theta_l, \theta'_h) = H(\theta'_h)$ , since the fact that  $\tilde{v}(\theta_h) > v^{-n}(\theta_h)$  implies that  $\tilde{\kappa} = H(\theta_h)$  (see in particular footnote 24). Hence,  $\tilde{\kappa}$  is differentiable in its second argument. Note also that as  $\tilde{\kappa}$  varies,  $\tilde{v}(\theta_l)$  remains fixed at  $v^{-n}(\theta_l)$  (one is not slack at both ends) and so  $\tilde{s}(\theta_l, \theta'_h)$  is feasible in  $\mathcal{P}(\theta_l, \theta_h)$ . But then, since  $\tilde{s}(\theta_l, \theta'_h)$  is optimal in  $\mathcal{P}(\theta_l, \theta_h)$  for all  $\theta'_h$  on a neighborhood of  $\theta_h$ , we have by what is essentially the envelope theorem that  $\int_{\theta_l}^{\theta_h} (\pi(\theta, \tilde{s}(\theta_l, \theta'_h)))_{\theta'_h} h(\theta) d\theta$  is well defined and equal to 0 evaluated at  $\theta_{h'} = \theta_h$ . Hence,

$$\begin{aligned} r_{\theta_h}(\theta_l, \theta_h) &= \left( \int_{\theta_l}^{\theta_h} \pi(\theta, \tilde{s}(\theta_l, \theta_h)) h(\theta) d\theta \right)_{\theta_h} \\ &= \pi(\theta_h, \tilde{s}(\theta_l, \theta_h)) h(\theta) + \int_{\theta_l}^{\theta_h} (\pi(\theta, \tilde{s}(\theta_l, \theta_h)))_{\theta_h} h(\theta) d\theta \\ &= \pi(\theta_h, \tilde{s}(\theta_l, \theta_h)) h(\theta). \end{aligned} \tag{11}$$

But since  $\tilde{\kappa}(\theta_l, \theta_h) = H(\theta_h)$ , and since  $(\theta_l, \theta_h)$  is a maximum of  $r$  and so  $r(\theta_l, \theta_h) > 0$  by Cn2, we have  $\pi(\theta_h, \tilde{s}(\theta_l, \theta_h)) > 0$  by Corollary 2 and, thus,  $r_{\theta_h}(\theta_l, \theta_h) > 0$ . Since  $(\theta_l, \theta_h)$  is optimal, it must thus be that  $\theta_h = 1$ . Similarly, if  $\tilde{v}(\theta_l) > v^{-n}(\theta_l)$ , then  $\theta_l = 0$ . But then, in all cases,  $\tilde{s}(\theta_l, \theta_h)$  is single-dominant on  $(\theta_l, \theta_h)$ , using stacking and Cn1. Part (iii) follows immediately, with the equality of payoffs following as the relevant domains of integration agree. Q.E.D.

LEMMA 8: *There exists  $(\underline{m}, \overline{m}) \ni \theta_\times^n$  such that  $\pi(\cdot, a^{-n}, v^{-n})$  is strictly positive on  $(\underline{m}, \overline{m})$ , strictly negative and strictly increasing for  $\theta < \underline{m}$ , and strictly negative and strictly decreasing for  $\theta > \overline{m}$ .*

PROOF: This follows from (6) since by Cn1 and stacking,  $a^{-n}$  is first strictly below  $n$ 's efficient action level and then strictly above, and so profits to imitation are single-peaked at  $\theta_\times^n$ . Formally, by stacking and Cn1, for  $\theta > \theta_\times^n$ ,  $a^{-n}(\theta) > \gamma(\theta, 0) \geq \gamma(\theta, H(\theta)) = \alpha_*(\theta)$ , and so  $\pi_a(\theta, a^{-n}, v^{-n}) < 0$ . Hence, anywhere that  $a^{-n}$  is differentiable, we have by (6) that  $(\pi(\theta, a^{-n}, v^{-n}))_\theta < 0$ . Further, at any point where  $a^{-n}$  jumps, say from  $a_l$  to  $a_h$ , we have, since  $v^{-n}$  is continuous and since  $a_h > a_l > \alpha_*(\theta)$ , that  $\pi(\theta, a_h, v^{-n}) - \pi(\theta, a_l, v^{-n}) < 0$ . Hence,  $\pi(\cdot, a^{-n}, v^{-n})$  is strictly decreasing on  $[\theta_\times^n, 1]$  and so single-crosses 0 from above at most once on  $[\theta_\times^n, 1]$ . If such a crossing exists, define  $\overline{m}$  as the crossing. If  $\pi(1, a^{-n}, v^{-n}) > 0$ , take  $\overline{m} = 1$ , and if  $\pi(\theta_\times^n, a^{-n}, v^{-n}) < 0$ , take  $\overline{m} = \theta_\times^n$ . Construct  $\underline{m}$  similarly. Q.E.D.

Strategy  $s^n$  is dominant on  $(\tau_l, \tau_h)$  if  $(\tau_l, \tau_h)$  is a maximal interval such that  $v^n > v^{-n}$ .

LEMMA 9: *Let  $(\alpha, v)$  be any feasible menu for  $n$ , let  $v$  be dominant on  $(\tau_l, \tau_h)$ , and let  $\pi(\cdot, \alpha, v) \geq 0$  on  $(\tau_l, \tau_h)$ . Then  $(\tau_l, \tau_h) \cap [\underline{m}, \overline{m}] \neq \emptyset$ .*

PROOF: Let  $\tau_l \geq \overline{m} \geq \theta_\times^n$ , where the case  $\tau_h \leq \underline{m}$  is similar. We will show that since the firm loses money with  $a^{-n}$  and  $v^{-n}$ , it a fortiori loses money with menu items that

implement an even more inefficiently high action and offer even more surplus. Note first that  $v(\tau_l) = v^{-n}(\tau_l)$  by definition of dominance, and since  $v$  and  $v^{-n}$  are continuous. Since for all  $\theta \in (\tau_l, \tau_h)$ ,

$$v(\tau_l) + \int_{\tau_l}^{\theta} \alpha(\tau) d\tau = v(\theta) > v^{-n}(\theta) = v^{-n}(\tau_l) + \int_{\tau_l}^{\theta} a^{-n}(\tau) d\tau,$$

it thus follows that there is  $\tau \in (\tau_l, \tau_h)$ , where  $\alpha(\tau) > a^{-n}(\tau)$ . But since  $\tau > \bar{m} \geq \theta_x^*$  and using Cn1, it follows that  $a^{-n}(\tau) > \alpha_x(\tau)$ , and so

$$\pi(\tau, \alpha(\tau), v(\tau)) < \pi(\tau, a^{-n}(\tau), v(\tau)) < \pi(\tau, a^{-n}(\tau), v^{-n}(\tau)) < 0,$$

a contradiction, since  $\pi(\theta, \alpha, v) \geq 0$  everywhere by hypothesis.

*Q.E.D.*

LEMMA 10: *Assume stacking. Fix  $n$  and  $s^{-n}$  satisfying Cn1 and Cn2. Then, for each  $\hat{s}$ , there is  $(\theta_l, \theta_h)$  with  $\Pi(\hat{s}, s^{-n}) \leq r(\theta_l, \theta_h)$ .*

PROOF: Intuitively, let  $\bar{m}^* \geq \bar{m}$  capture any region of dominance of  $v$  that contains  $\bar{m}$  and let  $\underline{m}^* \leq \underline{m}$  similarly capture any region of dominance of  $v$  that contains  $\underline{m}$ . Relative to  $\hat{s}$ , we will show that the firm strictly benefits by removing any agent it is winning outside of  $[\underline{m}^*, \bar{m}^*]$  and adding any agent in  $(\underline{m}, \bar{m})$  that it does not already serve with probability 1. But  $\bar{s}(\underline{m}^*, \bar{m}^*)$  accomplishes exactly this and does so optimally in the relaxed problem, and, hence, its associated payoff  $r(\underline{m}^*, \bar{m}^*)$  is at least as high as  $\Pi(s^n, s^{-n})$ .

To formalize this, note that using Proposition 5, we can without loss of generality (wlog) assume that  $(\alpha, v)$  loses money nowhere. Recall that  $\Pi(s) = \int_0^1 \pi(\theta, \alpha, v)\varphi(\theta, s)h(\theta) d\theta$ . Assume that  $v$  dominates  $v^{-n}$  on an interval  $I_H$  with  $\theta^x \leq \underline{I}_H \leq \bar{m} \leq \bar{I}_H$ . In this case, define  $\bar{m}^* = \bar{I}_H$ . If there is no such interval, define  $\bar{m}^* = \bar{m}$ . Similarly, if  $v$  dominates  $v^{-n}$  on an interval  $I_L$  with  $\underline{I}_L \leq \underline{m} \leq \underline{I}_L \leq \theta_x^*$ , then define  $\underline{m}^* = \underline{I}_L$ , and if there is no such interval, define  $\underline{m}^* = \underline{m}$ .

Consider first any positive-length interval  $J \subseteq [\bar{m}^*, 1]$  on which  $v = v^{-n}$  and such that  $\int_J \varphi(\theta, s) d\theta > 0$ . Then  $\alpha = a^{-n}$  on this interval and so, since  $\bar{m}^* \geq \bar{m}$ ,  $\pi(\theta, \alpha, v) < 0$  for all  $\theta > \bar{m}^*$ . Hence, excluding  $J$  from the domain of the integral in  $\Pi$  increases its value.

By Lemma 9 and since we have wlog taken  $(\alpha, v)$  to strictly lose money nowhere, there is no positive-length interval  $J = (\underline{J}, \bar{J})$  with  $\underline{J} \geq \bar{m}^*$  or and  $\bar{J} < \underline{m}^*$  on which  $v$  is dominant. We thus have  $\Pi(s) \leq \int_{\underline{m}^*}^{\bar{m}^*} \pi(\theta, \alpha, v)\varphi(\theta, s)h(\theta) d\theta$ . Define  $\hat{v} = \max(v, v^{-n})$ , with associated  $\hat{\alpha}$ , where at all  $\theta$  where  $v(\theta) \geq v^{-n}(\theta)$ , we can take  $\hat{\alpha} = \alpha$ , and at almost all  $\theta$  where  $v(\theta) \leq v^{-n}(\theta)$ , we can take  $\hat{\alpha} = a^{-n}$  (on any interval where  $v(\theta) = v^{-n}(\theta)$ ,  $\alpha = a^{-n}$  almost everywhere, and so there is a zero measure set where the two definitions might be in conflict). But then everywhere that  $\varphi(\theta, s)$  is positive (and so  $v(\theta) \geq v^{-n}(\theta)$ ), we have  $\pi(\theta, \hat{\alpha}, \hat{v}) = \pi(\theta, \alpha, v)$  and so  $\Pi(s) \leq \int_{\underline{m}^*}^{\bar{m}^*} \pi(\theta, \hat{\alpha}, \hat{v})\varphi(\theta, s)h(\theta) d\theta$ .

Consider any  $\theta \in (\underline{m}^*, \bar{m}^*)$  such that  $\varphi(\theta, s) < 1$ . Since by construction,  $\varphi$  is 1 on  $I_H$  and  $I_L$  (if these sets exist), it follows that  $\theta \in [\underline{m}, \bar{m}]$ . Then  $v(\theta) \leq v^{-n}(\theta)$  and so  $\hat{v}(\theta) = v^{-n}(\theta)$ , and  $\hat{\alpha}(\theta) = a^{-n}(\theta)$  almost everywhere and, thus,  $\pi(\theta, \hat{\alpha}, \hat{v}) = \pi(\theta, a^{-n}, v^{-n}) \geq 0$ . We thus have  $\Pi(s) \leq \int_{\underline{m}^*}^{\bar{m}^*} \pi(\theta, \hat{\alpha}, \hat{v})h(\theta) d\theta$ . But  $\hat{v} \geq v^{-n}$  by construction, and so (2) and (3) are satisfied in  $\mathcal{P}(\underline{m}^*, \bar{m}^*)$ , while  $\hat{\alpha}$  was chosen to be a subgradient of the convex function  $\max(v, v^{-n})$ , and, hence, (4) holds as well. Thus,  $(\hat{\alpha}, \hat{v})$  is feasible in  $\mathcal{P}(\underline{m}^*, \bar{m}^*)$ , from which  $\Pi(s) \leq r(\underline{m}^*, \bar{m}^*)$ .

*Q.E.D.*

PROOF OF PROPOSITION 1: The proof is immediate from Lemmas 7 and 10, as discussed in the main text.

*Q.E.D.*

C.2. Proofs for Section 6.1

LEMMA 11: Any maximum of  $r$  is in  $R = [0, \theta_x^n] \times [\theta_x^n, 1]$ .

PROOF: Let  $(\theta_l, \theta_h)$  with  $\theta_h < \theta_x^n$  be a maximum of  $r$ . Then  $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$  by Lemma 7, and so, since  $\tilde{\kappa} \leq H(\theta_h) < H(\theta_x^n) \leq 1$ , it follows from stacking and the definition of  $\theta_x^n$  that  $a^{-n} < \gamma(\cdot, \tilde{\kappa})$  for  $\theta < \theta_x^n$ , and so  $v$  crosses  $v^{-n}$  from below at  $\theta_h$ . But by Lemma 7(iii),  $v$  is single-dominant on  $(\theta_l, \theta_h)$ , a contradiction. Thus,  $\theta_h \geq \theta_x^n$ . Similarly,  $\theta_l \leq \theta_x^n$ . Q.E.D.

C.2.1. Local Properties of  $r$

For given function  $f$ , write  $f_x^+$  and  $f_x^-$  for the right and left derivatives of  $f$  with respect to  $x$ .

LEMMA 12: Considered as a function on  $\tilde{R}$ , that  $r$  is continuously differentiable, with

$$r_{\theta_h}(\theta_l, \theta_h) = (\pi_a(\theta_h, \gamma(\theta_h, \tilde{\kappa}), \tilde{v})(a^{-n}(\theta_h) - \gamma(\theta_h, \tilde{\kappa})) + \pi(\theta_h, \gamma(\theta_h, \tilde{\kappa}), \tilde{v}))h(\theta_h), \tag{12}$$

$$r_{\theta_l}(\theta_l, \theta_h) = (\pi_a(\theta_l, \gamma(\theta_l, \tilde{\kappa}), \tilde{v})(\gamma(\theta_l, \tilde{\kappa}) - a^{-n}(\theta_l)) - \pi(\theta_l, \gamma(\theta_l, \tilde{\kappa}), \tilde{v}))h(\theta_l). \tag{13}$$

PROOF: The right side of (12) has the same form as (5). As in the analysis of OB in the Supplemental Material Section S1.4, this is the value of increasing  $\theta_h$  by increasing the action immediately to the left of  $\theta_h$ , and since  $\gamma(\cdot, \tilde{\kappa}(\theta_l, \theta_h))$  solves the relaxed problem, this is as good as anything, and similarly for (13).<sup>48</sup> On  $\tilde{R}$ ,  $r_{\theta_h}$  and  $r_{\theta_l}$  are continuous. Hence,  $r$  is continuously differentiable. Q.E.D.

As a coherence check, along the lower boundary of  $\tilde{R}$  (and similarly on other boundaries),

$$r_{\theta_h}^+(\theta_l, \theta_h) = \lim_{\varepsilon \downarrow 0} \frac{r(\theta_l, \theta_h + \varepsilon) - r(\theta_l, \theta_h)}{\varepsilon} = \lim_{\varepsilon \downarrow 0} r_{\theta_h}(\theta_l, \theta_h + \varepsilon) = (r|_{\tilde{R}})_{\theta_h}(\theta_l, \theta_h), \tag{14}$$

where the second equality uses l’Hospitol’s rule and the third uses continuity of  $r_{\theta_h}$  on  $(\iota_l', \iota_h')$ .

Recall that  $\Theta$  is the subset of  $R$  on which  $\iota(\theta_l, \theta_h, \tilde{\kappa}(\theta_l, \theta_h)) = 0$ . Where there is no ambiguity, we write  $\tilde{\kappa}$  for  $\tilde{\kappa}(\theta_l, \theta_h)$ .

LEMMA 13: Consider  $r$  as a function on  $\tilde{R} \cap \Theta$ . Then  $r_{\theta_l \theta_h} < 0$ . If  $r_{\theta_h}(\theta_l, \theta_h) = 0$ , then  $r_{\theta_h \theta_h}(\theta_l, \theta_h) < 0$ , if  $r_{\theta_l}(\theta_l, \theta_h) = 0$ , then  $r_{\theta_l \theta_l}(\theta_l, \theta_h) < 0$ , and if  $r_{\theta_l}(\theta_l, \theta_h) = r_{\theta_h}(\theta_l, \theta_h) = 0$ , then  $r$  is locally strictly concave at  $(\theta_l, \theta_h)$ .

This proof differentiates (12) and (13), and uses that  $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$ , and hence  $(\tilde{v}(\theta_h))_{\theta_h} = (v^{-n}(\theta_h))_{\theta_h} = a^{-n}(\theta_h)$  and similarly at  $\theta_l$ . See Supplemental Material Section 6 for details.

---

<sup>48</sup>Alternatively, taking  $\tilde{s}(\theta_l, \theta_h)$  to have action profile  $\gamma(\cdot, \tilde{\kappa}(\theta_l, \theta_h))$  and the associated surplus function, then  $r(\theta_l, \theta_h) = \int_{\theta_l}^{\theta_h} \pi(\theta, \tilde{s}(\theta_l, \theta_h))h(\theta) d\theta$ . Integrate the surplus function out of this expression in the standard way (see Lemma 22 in the Supplemental Material) and then differentiate with respect to  $\theta_h$  and manipulate.

C.2.2. *Essentially Unique Optimality*

Fix a function  $f : [x_l, x_h] \rightarrow \mathbb{R}$ , where  $f$  is continuous, and has well defined and almost everywhere continuous one-sided derivatives. Say that  $x \in (x_l, x_h)$  is a critical point of  $f$  if  $f_x^-(x)f_x^+(x) \leq 0$ , so that  $f_x$  at least weakly changes sign at  $x$ . This includes the case where  $f$  is differentiable at  $x$  and  $f_x(x) = 0$ . Say that  $x_l$  is a critical point of  $f$  if  $f_x(x_l) \equiv f_x^+(x_l) \leq 0$  and that  $x_h$  is a critical point of  $f$  if  $f_x(x_h) \equiv f_x^-(x_h) \geq 0$ . Any maximum of  $f$  is at a critical point.

Note that  $\iota$  is continuously differentiable on each  $\tilde{R}$ , and is continuous on  $R$ . Hence, since  $\iota_{\theta_l}$ ,  $\iota_{\theta_h}$ , and  $\iota_\kappa$  are strictly positive, and since  $\iota(\theta_x^n, \theta_x^n, \cdot) = 0$ ,  $L_N$  is continuous, strictly decreasing, and goes through  $(\theta_x^n, \theta_x^n)$ . The locus  $L_S < L_N$  has the same properties.

LEMMA 14: *On or below  $L_S$ ,  $r_{\theta_h}(\theta_l, \theta_h) = \pi(\theta_h, \tilde{s}(\theta_l, \theta_h))h(\theta_h)$  and if  $r(\theta_l, \theta_h) > 0$ , then  $r_{\theta_h}(\theta_l, \theta_h) > 0$ . On or above  $L_N$ ,  $r_{\theta_l}(\theta_l, \theta_h) = -\pi(\theta_l, \tilde{s}(\theta_l, \theta_h))h(\theta_l)$ , and if  $r(\theta_l, \theta_h) > 0$ , then  $r_{\theta_l}(\theta_l, \theta_h) < 0$ .*

PROOF: Fix  $(\theta_l, \theta_h)$  below  $L_S$ . Then  $\iota(\theta_l, \theta_h, H(\theta_h)) < 0$ , and so by definition,  $\tilde{\kappa}(\theta_l, \theta_h) = H(\theta_h)$ , and by Lemma 1,  $\tilde{v}(\theta_l) = v^{-n}(\theta_l)$ , and, thus,  $\tilde{v}(\theta_h) > v^{-n}(\theta_h)$ . But then, as in (11),  $r_{\theta_h}(\theta_l, \theta_h) = \pi(\theta_h, \tilde{s}(\theta_l, \theta_h))h(\theta_h)$ . If  $r(\theta_l, \theta_h) > 0$ , then, since  $\tilde{\kappa} = H(\theta_h)$ , we have by Corollary 2 that  $\pi(\theta_h, \tilde{s}(\theta_l, \theta_h)) > 0$  and, hence,  $r_{\theta_h}(\theta_l, \theta_h) > 0$ .

Consider next  $(\theta_l, \theta_h) \in L_S$ . Since for each  $\varepsilon > 0$ ,  $(\theta_l, \theta_h - \varepsilon)$  is below  $L_S$ ,  $r_{\theta_h}(\theta_l, \theta_h - \varepsilon) = \pi(\theta_h - \varepsilon, \tilde{s}(\theta_l, \theta_h - \varepsilon))h(\theta_h - \varepsilon)$  by the previous step. It thus follows as in (14) that  $r_{\theta_h}^-(\theta_l, \theta_h) = \pi(\theta_h, \tilde{s}(\theta_l, \theta_h))h(\theta_h)$ , where we note that on  $L_S$ ,  $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$ . Finally, from (12) and the discussion immediately following Lemma 12, and again exploiting that above  $L_S$ ,  $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$ ,

$$\begin{aligned} r_{\theta_h}^+(\theta_l, \theta_h) &= \lim_{\varepsilon \downarrow 0} r_{\theta_h}(\theta_l, \theta_h + \varepsilon) \\ &= \lim_{\varepsilon \downarrow 0} \pi_a(\theta_h + \varepsilon, \tilde{s}(\theta_l, \theta_h + \varepsilon)) \\ &\quad \times (a^{-n}(\theta_h + \varepsilon) - \gamma(\theta_h + \varepsilon, H(\theta_h + \varepsilon)))h(\theta_h + \varepsilon) \\ &\quad + \lim_{\varepsilon \downarrow 0} \pi(\theta_h + \varepsilon, \tilde{s}(\theta_l, \theta_h + \varepsilon))h(\theta_h + \varepsilon) \\ &= \pi(\theta_h, \tilde{s}(\theta_l, \theta_h))h(\theta_h), \end{aligned}$$

where this follows since  $a^{-n}(\cdot) - \gamma(\cdot, H(\theta_h))$  is bounded and since  $\lim_{\varepsilon \downarrow 0} \pi_a(\theta_h + \varepsilon, \tilde{s}(\theta_l, \theta_h + \varepsilon)) = 0$  using that  $\gamma$  and  $v^{-n}$  are continuous, and that on  $L_S$ ,  $\tilde{\kappa} = H(\theta_h)$ , and, hence,  $\pi_a(\theta_h, \tilde{s}(\theta_l, \theta_h)) = 0$  by definition of  $\gamma$ . But then  $r_{\theta_h}^+(\theta_l, \theta_h) = r_{\theta_h}^-(\theta_l, \theta_h)$ , and so  $r_{\theta_h}(\theta_l, \theta_h)$  exists and has the claimed value. The proof for  $(\theta_l, \theta_h)$  above  $L_N$  is similar. *Q.E.D.*

ASSUMPTION 1— $L_S$  hits the western boundary of  $R$ : *We have  $\iota(0, 1, 1) \geq 0$ .*

Note in particular that since  $\iota$  is increasing, Assumption 1 implies that  $\iota(\theta_l, 1, 1) > 0$  for all  $\theta_l > 0$ , so that  $L_S$  does not intersect with the northern boundary of  $R$ .

Define  $\theta_T$  by  $\iota(0, \theta_T, 0) = 0$  if there is such a  $\theta_T \geq \theta_x^n$ , and by  $\theta_T = 1$  otherwise. This is the latitude at which  $L_N$  exits  $R$ . For firm 1,  $\theta_x^1 = 0$  and, hence,  $\theta_T = 0$ . Let  $A = \{(0, \theta_h) | \theta_h \geq \theta_T\}$ .

COROLLARY 3: Any maximum of  $r$  occurs in  $(\Theta \cup A) \setminus (\theta_x^n, \theta_x^n)$ .

PROOF: Since  $r$  is strictly positive at an optimum, then below  $L_S$ ,  $r_{\theta_h} > 0$  by Lemma 14, contradicting optimality, and above  $L_N$ ,  $r_{\theta_l} < 0$ , contradicting optimality unless  $\theta_l = 0$ . Hence, we must be in  $\Theta \cup A$ . But  $r(\theta_x^n, \theta_x^n) = 0$  and so  $(\theta_x^n, \theta_x^n)$  is not an optimum either. Q.E.D.

Note that  $\max_{\{(\theta_l, \theta_h) | \theta_h \geq \theta_l\}} r(\theta_l, \theta_h) = \max_{\Theta \cup A} r(\theta_l, \theta_h) = \max_{\theta_h} \psi(\theta_h)$ .

LEMMA 15: Fix  $\theta_h \in D$ . Then on  $\Theta(\theta_h)$ ,  $r(\cdot, \theta_h)$  is strictly single-peaked and has a unique maximum  $\lambda(\theta_h)$ . The function  $\psi$  is continuous on  $[\theta_x^n, 1]$ . On  $D$ ,  $\lambda$  is continuous as well.

PROOF: This is trivial for  $\theta_h > \theta_T$ , since  $\Theta(\theta_h) = \{0\}$ . Fix  $\theta_h \leq \theta_T$ . Let the (closed) interval  $\Theta(\theta_h)$  be denoted  $[\tau_l, \tau_h]$ . Existence of a maximum follows since  $r(\cdot, \theta_h)$  is continuous. Consider  $\theta_l \in [\tau_l, \tau_h]$ . If  $\theta_l \notin K$  and  $\theta_l$  is a critical point, then  $r_{\theta_l} = 0$  and so by Lemma 13,  $r_{\theta_l} < 0$ . Thus,  $\theta_l$  is a strict local maximum.

To show that  $r(\cdot, \theta_h)$  is strictly single-peaked, we will show that any critical point of  $r(\cdot, \theta_h)$  is a strict local maximum. Assume  $\theta_l \in K$  and that  $\theta_l$  is a critical point. Then, since  $\pi_a(\theta_l, \gamma(\theta_l, \bar{\kappa}), \tilde{v}) \geq 0$  and since  $a^{-n}$  jumps upward at  $\theta_l$ , we have by (13) that  $r_{\theta_l}^-(\theta_l, \theta_h) \geq r_{\theta_l}^+(\theta_l, \theta_h)$ . If  $r_{\theta_l}^- < 0$ , then  $r_{\theta_l}^+ < 0$ , contradicting that  $\theta_l$  is a critical point. Thus,  $r_{\theta_l}^- \geq 0$ . If  $r_{\theta_l}^- > 0$ , then  $r(\theta_l, \theta_h) > r(\theta'_l, \theta_h)$  for all  $\theta'_l$  in a neighborhood to the left of  $\theta_l$ . If instead  $r_{\theta_l}^- = 0$ , then, by 13 applied to the rectangle  $\tilde{R}$  to the left of  $(\theta_l, \theta_h)$ ,  $r(\cdot, \theta_h)$  is strictly concave on a neighborhood to the left of  $\theta_l$ , and so, since  $(r|_{\tilde{R}})_{\theta_l} = r_{\theta_l}^- = 0$ ,  $r(\cdot, \theta_h)$  is strictly increasing on that neighborhood. Thus, again,  $r(\theta_l, \theta_h) > r(\theta'_l, \theta_h)$  for all  $\theta'_l$  in a neighborhood to the left of  $\theta_l$ . Arguing similarly, if  $r_{\theta_l}^+(\theta_l, \theta_h) > 0$ , then we do not have a critical point, if  $r_{\theta_l}^+(\theta_l, \theta_h) = 0$  then  $r(\cdot, \theta_h)$  is strictly concave on a neighborhood to the right of  $\theta_l$  and so strictly decreasing on that neighborhood, and, finally, if  $r_{\theta_l}^+(\theta_l, \theta_h) < 0$ , then  $r(\cdot, \theta_h)$  is again strictly decreasing on a neighborhood to the right of  $\theta_l$ . It follows that  $r(\cdot, \theta_h)$  is strictly single-peaked on  $[\tau_l, \tau_h]$  and, hence, has a single optimum on  $[\tau_l, \tau_h]$ .

Since  $L_N$  and  $L_S$  are strictly decreasing and continuous, with  $L_N$  above  $L_S$ , the correspondence  $\Theta(\cdot)$  is nonempty, compact-valued, and continuous, and so by the theorem of the maximum,  $\psi$  is continuous and the set of maximizers of  $r(\cdot, \theta_h)$  is upper hemicontinuous in  $\theta_h$ . But then, since  $\lambda$  is single-valued on  $D$  by the first part of the proof, it is continuous as a function on  $D$ . Q.E.D.

LEMMA 16: The set  $D = \{\theta_h > \theta_x^n | \psi(\theta_h) > 0\}$  is an interval of the form  $(\theta_x^n, \bar{D})$ .

PROOF: Let  $\hat{\theta}_h \in D$ , so that  $r(\hat{\theta}_l, \hat{\theta}_h) > 0$  for some  $\hat{\theta}_l$ . We will show that  $(\theta_x^n, \hat{\theta}_h) \subseteq D$ . Let  $\hat{s} = (\hat{\alpha}, \hat{v})$  be the optimal strategy in  $\mathcal{P}(\hat{\theta}_l, \hat{\theta}_h)$ . We will show that for any  $\theta'_h \in (\theta_x^n, \hat{\theta}_h)$ , a parallel shift downward of  $\hat{v}$  wins a positive interval of types  $(\theta'_l, \theta'_h)$  strictly profitably and, hence,  $\psi(\theta'_h) > 0$ . Note first that PP (and in particular Proposition 5 in the Supplemental Material) implies that without loss of generality,  $\pi(\cdot, \hat{s}) \geq 0$ . Consider  $s' = (\hat{\alpha}, \hat{v} - \delta)$ , where  $\delta = \hat{v}(\theta'_h) - v^{-n}(\theta'_h)$ . By Cn1 and stacking,  $\hat{v}$  is strictly shallower than  $v^{-n}$  above  $\theta_x^n$ . So  $\delta > 0$ , and since  $\hat{v} - \delta$  is strictly steeper than  $v^{-n}$  below  $\theta_x^n$ , it follows that  $s'$  is single-dominant on  $(\theta'_l, \theta'_h)$  for some  $\theta'_l < \theta_x^n$ , and so a fortiori is feasible in  $\mathcal{P}(\theta'_l, \theta'_h)$ . But then, since  $\pi(\cdot, \hat{s}) \geq 0$ ,  $\pi(\cdot, s') \geq \delta > 0$  on  $(\theta'_l, \theta'_h) \supseteq (\theta_x^n, \theta'_h)$ . Hence,  $\psi(\theta'_h) \geq r(\theta'_l, \theta'_h) > 0$  and so  $\theta'_h \in D$ . Q.E.D.

LEMMA 17: Let  $(\lambda(\theta_h), \theta_h) \in L_S$  with  $\theta_h \in D$ . Then  $\psi$  is strictly increasing at  $\theta_h$  and so  $\theta_h$  is not a critical point of  $\psi$ .

PROOF: The basic idea is that by Lemma 14,  $r_{\theta_h}(\theta_l, \theta_h) > 0$  anywhere near  $L_S$ . But since  $L_S$  is decreasing, as one moves a little above  $L_S$ , the constraint on  $\theta_l$  is relaxed. Hence,  $\psi_{\theta_h}(\theta_h) \geq r_{\theta_h}(\theta_l, \theta_h) > 0$ . We need to account for the presence of kinks.

Let  $(\theta_l, \theta_h) = (\lambda(\theta_h), \theta_h) \in L_S$  with  $\theta_h \in D$  and, hence,  $\theta_h > \theta_h^n$ . Since  $K$  is finite, there is  $\delta > 0$  such that  $(\theta_h - \delta, \theta_h) \cap K = \emptyset$ , such that  $(\theta_l, \theta_l + \delta) \cap K = \emptyset$ , and, using continuity of  $\tilde{\kappa}$ , such that  $\tilde{\kappa}(\theta_l + \delta, \theta_h) > H(\theta_l)$ , so that all of  $\hat{X} = (\theta_l, \theta_l + \delta) \times (\theta_h - \delta, \theta_h)$  lies strictly below  $L_N$ . From Lemma 12,  $r_{\theta_h}$  is continuous on  $\hat{X}$ . Further, since  $\pi_a(\theta_h, \gamma(\cdot, \tilde{\kappa}(\theta_l, \theta_h)), \tilde{v}) = \pi_a(\theta_h, \gamma(\cdot, H(\theta_h)), \tilde{v}) = 0$  on  $L_S$ , by examination of (12) and by Lemma 14, it follows that  $r_{\theta_h}$  is continuous on  $X \equiv \hat{X} \cup \{(\theta_l, \theta_h)\}$ . Since  $\theta_h \in D$ ,  $r(\theta_l, \theta_h) > 0$ , and so by Lemma 14,  $r_{\theta_h}(\theta_l, \theta_h) > 0$ .

Note next that for each  $\theta'_h$  such that  $(\lambda(\theta'_h), \theta'_h) \in X \cap \Theta$ ,

$$\begin{aligned} \psi_{\theta'_h}^+(\theta'_h) &= \lim_{\varepsilon \downarrow 0} \frac{\psi(\theta'_h + \varepsilon) - \psi(\theta'_h)}{\varepsilon} \\ &\geq \lim_{\varepsilon \downarrow 0} \frac{r(\lambda(\theta'_h), \theta'_h + \varepsilon) - r(\lambda(\theta'_h), \theta'_h)}{\varepsilon} = r_{\theta_h}(\lambda(\theta'_h), \theta'_h), \end{aligned} \tag{15}$$

where the inequality follows since for small  $\varepsilon$ ,  $\lambda(\theta'_h)$  is feasible at  $\theta'_h + \varepsilon$ .<sup>49</sup> Thus, since  $r_{\theta_h}(\lambda(\theta_h), \theta_h) > 0$  by Lemma 14,  $\psi_{\theta_h}^+(\theta_h) > 0$ .

Finally, let us show that  $\psi_{\theta_h}^-(\theta_h) > 0$ . Let  $\rho = r_{\theta_h}(\lambda(\theta_h), \theta_h)/2$ . Since  $r_{\theta_h}$  is continuous on  $X$  and by (15), there is  $\hat{\varepsilon} > 0$  such that for all  $\tau \in [\theta_h - \hat{\varepsilon}, \theta_h]$ ,  $\psi_{\theta_h}^+(\tau) > \rho$ . To show that  $\psi_{\theta_h}^-(\theta_h) \geq \rho$ , it is sufficient that for any  $\varepsilon \in (0, \hat{\varepsilon})$ ,  $j(\theta_h) \geq 0$ , where for  $\tau \in [\theta_h - \varepsilon, \theta_h]$ ,

$$j(\tau) = \psi(\tau) - \psi(\theta_h - \varepsilon) - \rho(\tau - (\theta_h - \varepsilon)) =_s \frac{\psi(\tau) - \psi(\theta_h - \varepsilon)}{\tau - (\theta_h - \varepsilon)} - \rho.$$

Note that  $j(\theta_h - \varepsilon) = 0$ . But then, since  $j_{\tau}^+(\tau) = \psi_{\theta_h}^+(\tau) - \rho > 0$  for any  $\tau \in [\theta_h - \varepsilon, \theta_h]$ , if  $j(\tau) \geq 0$ , then  $j(\cdot) > 0$  for some interval to the right of  $\tau$  by the definition of a right derivative. Thus,  $j(\theta_h) \geq 0$ , and we are done. Q.E.D.

Let  $D' = \{\theta_h \in D \mid (\lambda(\theta_h), \theta_h) \notin L_S\}$  be the set of places where  $\lambda$  does not lie on  $L_S$ .

LEMMA 18: On  $D' \setminus K$ ,  $\psi$  is continuously differentiable, with  $\psi_{\theta_h}(\theta_h) = r_{\theta_h}(\lambda(\theta_h), \theta_h)$ . For all  $\theta_h \in D' \cap K$ ,  $\psi$  and  $r$  are right and left differentiable, with  $\psi_{\theta_h}^+(\theta_h) = r_{\theta_h}^+(\lambda(\theta_h), \theta_h)$  and  $\psi_{\theta_h}^-(\theta_h) = r_{\theta_h}^-(\lambda(\theta_h), \theta_h)$ .

PROOF: Let  $K_1 = (K \cap \lambda(D')) \cup \{0\}$  and  $K_2 = K \cap D'$ . In principle,  $r(\lambda(\cdot), \cdot)$  may be nondifferentiable because either  $\theta_h \in K_2$  or  $\lambda(\theta_h) \in K_1$ . There are thus several cases to consider.

**Case 1.** Consider first  $\theta_h \in D'$  such that  $\lambda(\theta_h) \notin K_1$  and  $\theta_h \notin K_2$ . We are not on  $L_S$  by definition of  $D'$  and we are not on  $L_N$  since by Lemma 14,  $r_{\theta_l}(\theta_l, \theta_h) < 0$  on  $L_N$ , contradicting the optimality of  $\lambda(\cdot)$ . Thus, since  $\psi(\theta_h) = r(\lambda(\theta_h), \theta_h)$  and since  $r$  is continuously

<sup>49</sup>That is,  $\lambda(\theta'_h) \in \Theta(\theta'_h + \varepsilon)$ , since  $L_S$  is decreasing and  $(\theta_l, \theta_l + \delta) \times (\theta_h - \delta, \theta_h)$  lies strictly below  $L_N$ .



differentiable on any given  $\tilde{R}$ ,  $\psi$  is also continuously differentiable, with

$$\psi_{\theta_h}(\theta_h) = r_{\theta_h}(\lambda(\theta_h), \theta_h) \tag{16}$$

since  $\lambda$  is continuous and by the envelope theorem.

**Case 2.** For any given  $\theta_l \in K_1$ , let  $J(\theta_l) = (\underline{J}(\theta_l), \bar{J}(\theta_l))$ , where  $\underline{J}(\theta_l) = \min\{\theta_h | \lambda(\theta_h) = \theta_l\}$  and  $\bar{J}(\theta_l) = \max\{\theta_h | \lambda(\theta_h) = \theta_l\}$ .<sup>50</sup> Since  $\lambda$  is constant on  $J(\theta_l)$ , if  $J(\theta_l)$  is nonempty, then for all  $\theta_h \in J(\theta_l) \setminus K_2$ , we have again have that  $\psi$  is continuously differentiable and (16).

**Case 3.** Consider next  $\theta_h \in (\{\underline{J}(\theta_l)\}_{\theta_l \in K_1} \cup \{\bar{J}(\theta_l)\}_{\theta_l \in K_1}) \setminus K_2$ . Assume that  $\theta_h = \underline{J}(\theta_l)$  for some  $\theta_l \in K_1$  (the case where  $\theta_h = \bar{J}(\theta_l)$  is similar). Then, for a neighborhood below  $\theta_h$ ,  $\theta'_h \notin K_2$ , since  $K_2$  is finite, and  $\lambda(\theta'_h) \notin K_1$  by definition of  $\underline{J}(\theta_l)$  and since  $K_1$  is finite. Hence,  $\psi_{\theta_h}(\theta'_h) = r_{\theta_h}(\lambda(\theta'_h), \theta'_h)$  by Case 1 and (16). If  $(\underline{J}(\theta_l), \bar{J}(\theta_l))$  is empty (that is, if  $\underline{J}(\theta_l) = \bar{J}(\theta_l)$ ), then by the exact same argument,  $\psi_{\theta_h}(\theta'_h) = r_{\theta_h}(\lambda(\theta'_h), \theta'_h)$  on a neighborhood immediately above  $\theta_h$ . Finally, if  $(\underline{J}(\theta_l), \bar{J}(\theta_l))$  is nonempty, then  $\psi_{\theta_h}(\theta'_h) = r_{\theta_h}(\lambda(\theta'_h), \theta'_h)$  for  $\theta'_h$  on a neighborhood above  $\theta_h$  by Case 2. Now note by (12) that  $r_{\theta_h}$  does not depend on  $a^{-n}(\theta_l)$ , and so  $r_{\theta_h}$  is continuous on  $D' \setminus K_2$ , even though  $\lambda(\theta_h) \in K_1$ .<sup>51</sup> But then, by continuity of  $\lambda$ ,  $\psi_{\theta_h}(\cdot)$  is continuously differentiable at  $\theta_h$ , again with  $\psi_{\theta_h}(\theta_h) = r_{\theta_h}(\lambda(\theta_h), \theta_h)$ .

**Case 4.** Finally, consider  $\theta_h \in K_2$ . Since  $K$  is finite, on some neighborhood above  $\theta_h$ ,  $\psi$  is continuously differentiable with  $\psi_{\theta_h} = r_{\theta_h}$  by the previous cases, and  $\lambda$  is continuous, and so

$$\begin{aligned} \psi_{\theta_h}^+(\theta_h) &= \lim_{\varepsilon \downarrow 0} \frac{\psi(\theta_h + \varepsilon) - \psi(\theta_h)}{\varepsilon} \\ &= \lim_{\varepsilon \downarrow 0} \psi_{\theta_h}(\theta_h + \varepsilon) \\ &= \lim_{\varepsilon \downarrow 0} r_{\theta_h}(\lambda(\theta_h + \varepsilon), \theta_h + \varepsilon) = r_{\theta_h}^+(\lambda(\theta_h), \theta_h), \end{aligned}$$

using l'Hospital's rule at the second inequality, and where the last equality uses from the last part of Case 3, that  $r_{\theta_h}$  is continuous on  $D' \setminus K_2$ . Similarly,  $\psi_{\theta_h}^-(\theta_h) = r_{\theta_h}^-(\lambda(\theta_h), \theta_h)$ . Q.E.D.

**LEMMA 19:** *Every critical point  $\theta_h$  of  $\psi$  on  $D$  is a strict local maximum of  $\psi$ . That is, for all  $\theta'_h \neq \theta_h$  in some neighborhood of  $\theta_h$ ,  $\psi(\theta'_h) < \psi(\theta_h)$ .*

**PROOF:** Assume first that  $\theta_h = 1$ . If  $\psi_{\theta_h}(1) > 0$ , then 1 is a strict local maximum. So, in what follows, assume that either  $\theta_h < 1$  or  $\theta_h = 1$ , with  $\psi_{\theta_h}(1) = 0$ . By Lemma 17, any critical point of  $\psi$  on  $D$  is an element of  $D'$ , except in the case that  $\theta_h = 1 \in D \setminus D'$ , in which case we are trivially done by Lemma 17. But on  $D'$ , Lemma 18 lets us relate the local concavity properties of  $\psi$  to those we establish for  $r$  in Lemma 13. We go through the same cases as in Lemma 18.

**Case 1.** Consider first a critical point  $\theta_h \in D'$  such that  $\lambda(\theta_h) \notin K_1$  and  $\theta_h \notin K_2$ . Then, since  $0 \in K_1$ ,  $(\lambda(\theta_h), \theta_h) \in \Theta$ , and so Lemma 13 applies and  $r_{\theta_l \theta_l} < 0$ , and, thus, by the

<sup>50</sup>These correspond to the bottoms and tops of the vertical segments of the path in Figure 5.

<sup>51</sup>Note that  $r_{\theta_h}$  depends on  $\theta_l$  only through  $\tilde{\kappa}$ , and that  $\tilde{\kappa}$  does not depend on  $a^{-n}(\theta_l)$ .

implicit function theorem,  $\lambda_{\theta_h} = -r_{\theta_l\theta_h}/r_{\theta_l\theta_l} < 0$ . Since (16) holds on a neighborhood of  $\theta_h$ , then

$$\begin{aligned} \psi_{\theta_h\theta_h}(\theta_h) &= r_{\theta_h\theta_l}(\lambda(\theta_h), \theta_h)\lambda_{\theta_h}(\theta_h) + r_{\theta_h\theta_h}(\lambda(\theta_h), \theta_h) \\ &= -\frac{(r_{\theta_l\theta_h}(\lambda(\theta_h), \theta_h))^2}{r_{\theta_l\theta_l}(\lambda(\theta_h), \theta_h)} + r_{\theta_h\theta_h}(\lambda(\theta_h), \theta_h) \\ &= \frac{1}{r_{\theta_l\theta_l}(\lambda(\theta_h), \theta_h)}(r_{\theta_l\theta_l}(\lambda(\theta_h), \theta_h)r_{\theta_h\theta_h}(\lambda(\theta_h), \theta_h) - (r_{\theta_l\theta_h}(\lambda(\theta_h), \theta_h))^2). \end{aligned}$$

Since  $\theta_h$  is a critical point,  $\psi_{\theta_h}(\theta_h) = 0$ , and so  $r_{\theta_l} = r_{\theta_h} = 0$  at  $(\lambda(\theta_h), \theta_h)$ . Thus, by Lemma 13  $r_{\theta_l\theta_l}r_{\theta_h\theta_h} - r_{\theta_l\theta_h}^2 > 0$ . Hence,  $\psi_{\theta_h\theta_h}(\theta_h) < 0$  and  $\theta_h$  is a strict local maximum of  $\psi$ .

**Case 2.** Consider  $\theta_h \in D'$ , where  $\theta_h \notin K_2$ , but for some  $\theta_l \in K_1$ ,  $\theta_h \in J(\theta_l)$ . Then, since  $J(\theta_l) \setminus K_2$  is open, by Case 2 of Lemma 18, (16) holds on a neighborhood of  $\theta_h$ , and so, since  $\lambda$  is constant on  $J(\theta_l)$ ,  $\psi(\cdot) = r(\theta_l, \cdot)$ , and so, for example,  $\psi_{\theta_h\theta_h}(\theta_h) = r_{\theta_h\theta_h}(\theta_l, \theta_h)$ . If  $\theta_h \leq \theta_T$ , so that  $(\lambda(\theta_h), \theta_h) \in \Theta$ , then by Lemma 13, if  $\psi_{\theta_h}(\theta_h) = 0$ , then  $\psi_{\theta_h\theta_h}(\theta_h) < 0$ , so  $\theta_h$  is a strict local maximum of  $\psi$ . Assume that  $\theta_h \geq \theta_T$ , so that  $\lambda(\theta_h) = 0$  and  $\tilde{\kappa}(\lambda(\theta_h), \theta_h) = 0$ . Trace the derivation of  $r_{\theta_h\theta_h}$  in the proof of Lemma 13 up through (26) with  $\tilde{\kappa}$  replaced by 0 and note that this part of the proof relies on  $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$  but not on  $\tilde{v}(\theta_l) = v^{-n}(\theta_l)$ . It follows that where  $r_{\theta_h}(0, \theta_h) = 0$ ,  $\psi_{\theta_h\theta_h}(0, \theta_h) = r_{\theta_h\theta_h}(\underline{\theta}, \theta_h) < 0$ , and again,  $\theta_h$  is a strict local maximum of  $\psi$ .

**Case 3.** Consider next  $\theta_h \in (\{J(\theta_l)\}_{\theta_l \in K_1} \cup \{\bar{J}(\theta_l)\}_{\theta_l \in K_1}) \setminus K$ . Assume that  $\theta_h = \underline{J}(\theta_l)$  for some  $\theta_l \in K_1$  (the other case is similar) and assume that  $\psi_{\theta_h}(\theta_h) = 0$ . Then by Case 2,  $\psi$  is strictly concave on a neighborhood just above  $\theta_h$ , while by Case 1,  $\psi$  is strictly concave on a neighborhood just below  $\theta_h$ . Hence, again,  $\theta_h$  is a strict local maximum of  $\psi$ .

**Case 4.** Finally, consider  $\theta_h \in K_2 = K \cap D'$ . Since  $\tilde{\kappa} \in [0, 1]$  and since  $\theta_h > \theta_x^n$ , we have that  $a^{-n} - \gamma$  is positive and bounded away from 0 and  $\infty$  on a neighborhood of  $\theta_h$  by stacking and Cn1. At any point  $\theta'_h$  of continuity of  $a^{-n}$ , and repeating (12) for convenience,

$$\begin{aligned} \frac{\psi_{\theta_h}(\theta'_h)}{h(\theta'_h)} &= \frac{r_{\theta_h}(\lambda(\theta'_h), \theta'_h)}{h(\theta'_h)} \\ &= \pi(\theta'_h, \gamma(\cdot, \tilde{\kappa}), v^{-n}(\theta'_h)) \\ &\quad + \pi_a(\theta'_h, \gamma(\cdot, \tilde{\kappa}), v^{-n}(\theta'_h))(a^{-n}(\theta'_h) - \gamma(\theta'_h, \tilde{\kappa})), \end{aligned} \tag{17}$$

where we recall that  $\tilde{\kappa}$  is continuous and, hence, so is  $\gamma(\cdot, \tilde{\kappa})$ , and that  $v^{-n}$  is also continuous and, hence, so are  $\pi$  and  $\pi_a$ . Thus, any discontinuity in  $\psi_{\theta_h}$  at  $\theta_h$  is driven by an upward jump of  $a^{-n}$  at  $\theta_h$  and, since  $\pi_a(\theta_h, \gamma(\cdot, \tilde{\kappa}), v^{-n}(\theta_h)) \leq 0$  (since  $\tilde{\kappa} \leq H(\theta_h)$ ), for there to be a discontinuity, we must have  $\pi_a < 0$ .

If  $\pi \leq 0$ , then, by (17), both  $\psi_{\theta_h}^+(\theta_h)$  and  $\psi_{\theta_h}^-(\theta_h)$  are strictly negative, and  $\theta_h$  is not a critical point. If  $\pi > 0$ , then since  $a^{-n}$  jumps up at  $\theta_h$ , we have  $\psi_{\theta_h}^- > \psi_{\theta_h}^+$ . Assume that  $\theta_h$  is a critical point, so that  $\psi_{\theta_h}^- \psi_{\theta_h}^+ \leq 0$ . If  $\psi_{\theta_h}^- > 0 > \psi_{\theta_h}^+$ , then  $\theta_h$  is a strict local maximum of  $\psi$ . If  $\psi_{\theta_h}^+ = 0$ , then, first,  $\psi_{\theta_h}^- > 0$ , and, second, from the previous cases,  $\psi_{\theta_h\theta_h} < 0$  for all  $\theta$  on a neighborhood to the right of  $\theta_h$ . Similarly if  $\psi_{\theta_h}^- = 0$ , then  $\psi_{\theta_h}^+ < 0$ , and  $\psi_{\theta_h\theta_h} < 0$  for all  $\theta$  on a neighborhood to the left of  $\theta_h$ . In each case  $\theta_h$  is again a strict local maximum of  $\psi$ . Q.E.D.

COROLLARY 4: *There is a unique critical point of  $\psi$  on  $[\theta_x^n, 1]$  and it uniquely maximizes  $\psi$ .*

PROOF: By Weierstrass' theorem,  $\psi$  has a maximum on  $[\theta_x^n, 1]$ . But for any maximizer,  $\theta_h^*$ ,  $\theta_h^* \neq \theta_x^n$ , since  $\Theta(\theta_x^n) = \{\theta_x^n\}$ , and so  $\psi(\theta_x^n) = r(\theta_x^n, \theta_x^n) = 0$ . If  $\theta_h^* \in (\theta_x^n, 1)$ , then  $\psi_{\theta_h}^-(\theta_h^*) \geq 0$ , and  $\psi_{\theta_h}^+(\theta_h^*) \leq 0$ , since  $\theta_h^*$  is a maximizer. Similarly, if  $\theta_h^* = 1$ , then  $\psi_{\theta_h}^-(\theta_h^*) \geq 0$ . In each case,  $\theta_h^*$  is critical by definition. Thus, since any maximum is a critical point,  $\psi$  has a critical point.

Let  $\theta_h^*$  be any critical point of  $\psi$  and let us show that  $\theta_h^*$  is the unique maximizer of  $\psi$  (and, hence,  $\theta_h^*$  is the unique critical point of  $\psi$ ). Without loss of generality, assume that for some  $\theta_h^{**} > \theta_h^*$ ,  $\psi(\theta_h^{**}) \geq \psi(\theta_h^*)$ . Then  $\psi$  attains a minimum  $\theta_h^{\min}$  on the compact set  $[\theta_h^*, \theta_h^{**}]$ . But since  $\theta_h^*$  is a strict local maximum of  $\psi$ , we have  $\psi(\theta_h^{\min}) < \psi(\theta_h^*) \leq \psi(\theta_h^{**})$ , and so  $\theta_h^{\min} \in (\theta_h^*, \theta_h^{**})$ . Hence, since  $\theta_h^{\min}$  is an interior minimum,  $\psi_{\theta_h}^-(\theta_h^{\min}) \leq 0$ , and  $\psi_{\theta_h}^+(\theta_h^{\min}) \geq 0$ , and so  $\theta_h^{\min}$  is a critical point of  $\psi$ . But then, by Lemma 19,  $\theta_h^{\min}$  is a strict local maximum, a contradiction. Q.E.D.

LEMMA 20: *Let  $\theta_h^*$  be the unique maximizer of  $\psi$ . Then the unique maximizer of  $r$  is  $(\lambda(\theta_h^*), \theta_h^*)$ .*

PROOF: Let  $(\theta_l^{**}, \theta_h^{**}) \in \arg \max_{\{(\theta_l, \theta_h) | 1 \geq \theta_h \geq \theta_l \geq 0\}} r(\theta_l, \theta_h)$ . Since  $D$  is nonempty,  $r(\theta_l^{**}, \theta_h^{**}) > 0$ , and, hence,  $\theta_l^{**} < \theta_h^{**}$  and  $\theta_h^{**} \in D$ . By Corollary 3,  $(\theta_l^{**}, \theta_h^{**}) \in \Theta \cup A$ , and so  $\theta_l^{**} \in \Theta(\theta_h^{**})$ . Hence, by Lemma 15,  $\theta_l^{**} = \lambda(\theta_h^{**})$ . Since  $(\theta_l^{**}, \theta_h^{**})$  is optimal and since the constraint  $\theta_h \geq \theta_l$  is slack, we must have  $r_{\theta_h}^+(\lambda(\theta_h^{**}), \theta_h^{**}) \leq 0$  and  $r_{\theta_h}^-(\lambda(\theta_h^{**}), \theta_h^{**}) \geq 0$ . But then, by Lemma 18,  $\psi_{\theta_h}^+(\theta_h^{**}) \leq 0$  and  $\psi_{\theta_h}^-(\theta_h^{**}) \geq 0$ , and so by Corollary 4,  $\theta_h^{**} = \theta_h^*$ , and we are done. Q.E.D.

### C.3. Proofs for Section 6.2

PROOF OF THEOREM 2: SUFFICIENCY: Let  $\hat{s}$  satisfy PS, IO, and OB. Fix  $n$  and let  $\hat{s}^n = (\hat{\alpha}, \hat{v})$ , with associated  $\hat{\kappa}$ . By IO,  $(\hat{\alpha}, \hat{v})$  satisfies Cn1 on  $[\theta_l, \theta_h]$ . But then, by IO, if  $n < N$ , then  $\pi_a(\theta_h, \hat{\alpha}, \hat{v}) < 0$ , and by Cn1 and stacking,  $a^{-n}(\theta_h) - \hat{\alpha}(\theta_h) > 0$ . Hence, by (5),  $\pi(\theta_h, \hat{\alpha}, \hat{v}) > 0$ . Similarly,  $\pi(\theta_l, \hat{\alpha}, \hat{v}) > 0$  if  $n > 1$ . But by Lemma 2, profits are strictly single-peaked with maximum at  $\theta_0$  solving  $H(\theta_0) = \hat{\kappa}$ , and so  $\pi(\theta, \hat{\alpha}, \hat{v}) > 0$  for all  $\theta \in [\theta_l, \theta_h]$ . Thus,  $\hat{v}(\theta) < v_*(\theta)$ , so that  $(\hat{\alpha}, \hat{v})$  satisfies Cn2 on  $[\theta_l, \theta_h]$ .

Let us redefine  $(\hat{\alpha}, \hat{v})$  outside of  $[\theta_l, \theta_h]$  to satisfy Cn1 and Cn2 there as well. Set  $\alpha(\theta)$  as  $\min\{\gamma(\theta, 0), \hat{\alpha}(\theta_l)\}$  for  $\theta < \theta_l$ ,  $\hat{\alpha}(\theta)$  for  $\theta \in [\theta_l, \theta_h]$ , and  $\max\{\gamma(\theta, 1), \hat{\alpha}(\theta_h)\}$  for  $\theta > \theta_h$ , and set  $v(\theta) = \hat{v}(\theta_l) + \int_{\theta_l}^{\theta} \alpha(\tau) d\tau$  for all  $\theta$ . That is, actions and surplus modified outside of  $[\theta_l, \theta_h]$  to ensure that Cn1 holds while respecting monotonicity. Note that  $\hat{\alpha}(\theta_h) = \gamma(\theta_h, \hat{\kappa}) \geq \gamma(\theta_h, 1)$  and so no discontinuity is introduced at  $\theta_h$ , and similarly at  $\theta_l$ . By stacking,  $(\alpha, v)$  is single-dominant on  $(\theta_l, \theta_h)$ , and so, since  $(\alpha, v)$  and  $(\hat{\alpha}, \hat{v})$  agree on  $[\theta_l, \theta_h]$ ,  $(\alpha, v)$  and  $(\hat{\alpha}, \hat{v})$  are equivalent.

To show that Cn2 holds for  $\theta \notin [\theta_l, \theta_h]$ , assume  $(\theta_h, 1)$  is nonempty (the case  $[0, \theta_l)$  nonempty is the same). Where  $\alpha(\cdot) = \hat{\alpha}(\theta_h)$ ,  $(\pi(\theta, \alpha, v))_{\theta} = \pi_a(\theta, \alpha, v)\alpha_{\theta}(\theta) = 0$  by (6). Where  $\alpha(\cdot) = \gamma(\cdot, 1)$ ,  $(\pi(\theta, \alpha, v))_{\theta} = \pi_a(\theta, \gamma(\cdot, 1), v)\gamma_{\theta}(\theta, 1) \geq 0$ , using that  $\gamma_{\theta}(\theta, 1) > 0$ , that  $\gamma(\theta, 1) \leq \gamma(\theta, H(\theta)) = \alpha_*(\theta)$ , and that  $\pi$  is strictly concave in  $a$ , and so  $\pi_a(\theta, \gamma(\cdot, 1), v) \geq 0$ . Thus,  $\pi(\theta, \alpha, v) \geq \pi(\theta_h, \alpha, v) > 0$  for all  $\theta > \theta_h$ , and so  $v(\theta) < v_*(\theta)$  and Cn2 holds everywhere.

Construct the strategy profile  $s$  by performing the above process for each  $n$ . Then OB continues to hold for all  $n$ , since for each of  $n$ 's opponents,  $\hat{\alpha}$  and  $\alpha$  agree on  $[\theta_l, \theta_h]$ ,

and since both the modified and original action profiles of  $n$ 's opponents are continuous. Let us show that  $s$  is a Nash equilibrium. Fix  $n \notin \{1, N\}$ . Assume first that Assumption 1 holds. By the first paragraph of this proof,  $\theta_h \in D$ . By PS,  $0 < \theta_l < \theta_h < 1$ , and so  $\iota(\theta_l, \theta_h, \tilde{\kappa}(\theta_l, \theta_h)) = 0$ , where  $\tilde{\kappa}(\theta_l, \theta_h) \in (H(\theta_l), H(\theta_h))$  by IO and OB (see the third paragraph after Lemma 2), and so  $\theta_l \in \Theta(\theta_h)$ . But then, since  $r_{\theta_l}(\theta_l, \theta_h) = 0$  by OB, we have  $\theta_l = \lambda(\theta_h)$  by Lemma 15. But then, again by OB,  $0 = r_{\theta_h}(\theta_l, \theta_h) = r_{\theta_h}(\lambda(\theta_h), \theta_h) = \psi_{\theta_h}(\theta_h)$ , where the third equality is by Lemma 18. Finally, since by Corollary 4,  $\psi$  is strictly single-peaked on the interval  $D$ ,  $\theta_h = \theta_h^*$  by Lemma 20. Thus,  $s^n$  is a best response to  $s^{-n}$  by Corollary 1.

If Assumption 1 fails, then recall from the end of Section 6.1 that  $\tilde{\lambda}$  is the analogue to  $\lambda$ . So we argue first that  $\theta_h \in \tilde{\Theta}(\theta_l)$ , second that, by the analogue to Lemma 15,  $\theta_h = \tilde{\lambda}(\theta_l)$ , and third that, by the analogue to Lemma 18 and by OB,  $0 = r_{\theta_l}(\theta_l, \theta_h) = r_{\theta_l}(\theta_l, \tilde{\lambda}(\theta_l)) = \tilde{\psi}_{\theta_l}(\theta_l)$ . But then, since  $\tilde{\psi}$  is strictly single-peaked on  $\tilde{D}$ , we have  $\theta_l = \tilde{\theta}_l^*$ , and again  $s^n$  is a best response to  $s^{-n}$ .

Consider  $n = 1$ . Then  $\kappa^1 = 0$  by IO and  $\theta_l = 0$  by PS. But since  $\kappa^1 = 0$  and since, by the first part of this proof,  $\pi(\theta_h, \hat{\alpha}, \hat{v}) > 0$ , by Lemma 2,  $\pi(\theta, \hat{\alpha}, \hat{v}) > 0$  for all  $\theta < \theta_h$ . Thus, since  $\pi_a(\theta, \hat{\alpha}, \hat{v}) < 0$  for all  $\theta < \theta_h$ ,  $r_{\theta_l} < 0$ , and so  $0 = \lambda(\theta_h)$ . But then, by OB,  $0 = r_{\theta_h}(0, \theta_h) = r_{\theta_h}(\lambda(\theta_h), \theta_h) = \psi_{\theta_h}(\theta_h)$ , and again  $\theta_h = \theta_h^*$ , and  $s^1$  is a best response to  $s^{-1}$ . Finally, consider  $n = N$ . Then  $\kappa^N = 1$  by IO, and so, as above,  $\theta_h = 1 = \tilde{\lambda}(\theta_l)$ . Thus, by OB,  $0 = r_{\theta_l}(\theta_l, 1) = r_{\theta_l}(\theta_l, \tilde{\lambda}(\theta_l)) = \tilde{\psi}_{\theta_l}(\theta_l)$ , and so  $\theta_l = \tilde{\theta}_l^*$ , and  $s^N$  is a best response to  $s^{-N}$ . Q.E.D.

**PROOF OF THEOREM 2: EXISTENCE:** We begin with three further restrictions on strategies that will not bind in equilibrium, but help us toward compactness and continuity. Recall that  $BR(s^{-n}) = \arg \max_{s^n \in S^n} \Pi^n(s^n, s^{-n})$ . Since  $\pi_{aa}^n = V_{aa}^n$ , by definition of  $\gamma$ , we have  $\gamma_\theta^n(\theta, \kappa) = (((\kappa - H(\theta))/h(\theta))_\theta - 1)/V_{aa}^n(\gamma^n(\theta, \kappa))$ . Since  $V_{aa}^n < 0$  is continuous, it is bounded away from zero on the set  $\{\gamma^n(\theta, \kappa) | \theta \in [0, 1], \kappa \in [0, 1]\}$ , which is compact since  $\gamma^n$  is continuous. But then, since  $h$  is  $C^1$  and bounded away from 0,  $b_1 = \max\{\gamma^N(1, 0), \max_{n, \theta \in [0, 1], \kappa \in [0, 1]} \gamma_\theta^n(\theta, \kappa)\}$  is well defined and finite. Since  $\gamma^n(\theta, \kappa) \leq \gamma^N(1, 0)$  for all  $n, \theta$ , and  $\kappa \in [0, 1]$ ,  $b_1$  is a bound on the highest value and slope of any  $\gamma$  satisfying Cn1. We will bound the slopes of actions profiles by  $b_1$ .

**Cn3.** The action function satisfies  $0 \leq \alpha^n(\theta') - \alpha^n(\theta) \leq b_1(\theta' - \theta)$  for all  $\theta, \theta'$  with  $\theta' > \theta$  and all  $n$ .

Next, let  $b_2 = \min_{n, \theta_h \in [0, 1], \kappa \in [0, 1]} (\pi_a^n(\theta_h, \gamma^n(\cdot, \kappa), 0)\gamma^N(\theta_h, 0) + \pi^n(\theta_h, \gamma^n(\cdot, \kappa), 0))$ , where  $b_2 > -\infty$  since each relevant object is continuous and, hence, bounded on the compact choice set. We will see that if  $v^n(1) < b_2$ , then for any  $\theta_h < 1$ , the incentive for firm  $n$  to grab market share by increasing  $\theta_h$  will be strictly positive, motivating our next restriction.

**Cn4.** The surplus function satisfies  $v^n(1) \geq b_2$  for all  $n$ .

For each  $n$ , let  $S_R^n$  be the subset of  $S^n$  such that conditions Cn1–Cn4 hold. Let  $S_R = \times_n S_R^n$  and  $S_R^{-n} = \times_{n' \neq n} S_{R}^{n'}$ . To see that  $S_R^n$  is nonempty, let us argue that  $(\alpha_*^n, v_*^n) \in S_R^n$ . Note that Cn2 is immediate and that Cn1 follows because  $\alpha_*^n(\theta) = \gamma^n(\theta, H(\theta))$ . But then, by definition of  $b_1$ ,

$$(\alpha_*^n(\theta))_\theta = \gamma_\theta^n(\theta, H(\theta)) + \gamma_\kappa^n(\theta, H(\theta))h(\theta) < \gamma_\theta^n(\theta, H(\theta)) \leq b_1,$$

using that  $\gamma_\kappa^n < 0$ , and so Cn3 follows. To see Cn4, note that since  $\alpha_*^n(1) = \gamma^n(1, 1)$ , it follows that  $\pi_a^n(1, \gamma^n(1, 1), 0) = 0$  and, hence,  $\pi_a^n(1, \gamma^n(1, 1), 0)\gamma^N(\theta_h, 0) + \pi^n(1, \gamma^n(1, 1), 0) = v_*^n(1)$ , noting that we are evaluating  $\pi^n$  at surplus to the agent of 0. Thus, since  $b_2$  is a minimum of objects of this form,  $v_*^n(1) \geq b_2$ . Hence,  $S_R^n$  is nonempty.

LEMMA 21: Fix  $s^{-n} \in S_R^{-n}$ . Then  $BR^n(s^{-n}) \cap S_R^n$  is nonempty.

PROOF: Fix  $n$  and fix  $\hat{s}^n \in BR^n(s^{-n})$ , where we note that  $BR^n(s^{-n})$  is nonempty since  $r$  has a maximizer and using Corollary 1. Further, by that corollary and using stacking,  $\hat{s}^n$  is single-dominant on some region  $(\theta_l, \theta_h)$  and has the form  $(\hat{\alpha}, \hat{v})$ , where  $\hat{\alpha} = \gamma(\cdot, \kappa)$  on  $[\theta_l, \theta_h]$ , where  $\kappa \in [H(\theta_l), H(\theta_h)]$ , and where Cn1 and Cn2 are satisfied on  $[\theta_l, \theta_h]$ . Let  $(\alpha, v)$  be defined from  $(\hat{\alpha}, \hat{v})$  as in the proof of Theorem 2, so that as shown there, Cn1 and Cn2 are satisfied on  $[0, 1]$ . By stacking and using that for  $n' \neq n$ , Cn1 and Cn2 are satisfied by assumption, it remains the case that  $(\alpha, v)$  is single-dominant on  $[\theta_l, \theta_h]$ , and since  $(\alpha, v)$  and  $(\hat{\alpha}, \hat{v})$  agree on  $[\theta_l, \theta_h]$ , it follows that  $(\alpha, v) \in BR(s^{-n})$ . Condition Cn3 holds by construction.

To show Cn4, assume by way of contradiction that  $v(1) < b_2$ . Then, since  $v^{n'}(1) \geq b_2$  for each  $n' \neq n$ , we have  $\theta_h < 1$ , and so by (12), if we let  $\bar{a} = \lim_{\theta'_h \downarrow \theta_h} a^{-n}(\theta'_h)$ , then, by Corollary 1, since  $(\theta_l, \theta_h)$  maximizes  $r$ ,  $0 \geq r_{\theta_h}^+(\theta_l, \theta_h) / h(\theta_h) = \pi_a(\theta_h, \gamma(\cdot, \kappa), v)(\bar{a} - \gamma(\theta_h, \kappa)) + \pi(\theta_h, \gamma(\cdot, \kappa), v)$ . But since  $s^n$  is a best response, Proposition 5 and continuity of  $\pi$ ,  $\gamma$ , and  $v$  yield that  $\pi(\theta_h, \gamma(\cdot, \kappa), v) \geq 0$ . By Cn1 and Cn2 for  $n' \neq n$  and by stacking,  $\bar{a} - \gamma(\theta_h, \kappa) > 0$ . Hence,  $0 \geq \pi_a(\theta_h, \gamma(\cdot, \kappa), v)$  and so

$$\begin{aligned} 0 &\geq \pi_a(\theta_h, \gamma(\cdot, \kappa), v) \gamma^N(\theta_h, 0) + \pi(\theta_h, \gamma(\cdot, \kappa), v) \\ &> \pi_a(\theta_h, \gamma(\cdot, \kappa), 0) \gamma^N(\theta_h, 0) + \pi(\theta_h, \gamma(\cdot, \kappa), 0) - b_2 \geq 0, \end{aligned}$$

where the first inequality uses  $\bar{a} - \gamma(\theta_h, \kappa) \leq \bar{a} \leq \gamma^N(\theta_h, 0)$ , the second uses monotonicity of  $v$  and  $v(1) < b_2$ , and the last uses the definition of  $b_2$ . This is a contradiction and, hence,  $v(1) \geq b_2$  as required. Since  $(\alpha, v)$  is a best response and satisfies Cn1–Cn4, we are done. Q.E.D.

Let us now prove that the game  $(S^n, \Pi^n)_{n=1}^N$  has a pure-strategy equilibrium. It is enough to show that  $(S_R^n, \Pi^n)_{n=1}^N$  has a pure-strategy equilibrium, since by Lemma 21,  $BR^n(s^{-n}) \cap S_R^n$  is nonempty, and so in a Nash equilibrium of  $(S_R^n, \Pi^n)_{n=1}^N$ , each player is playing an element of  $BR^n(s^{-n})$ , and we have a Nash equilibrium of  $(S^n, \Pi^n)_{n=1}^N$ .

The set of continuous functions from  $[0, 1]$  to  $\mathbb{R}$ , endowed with the sup norm  $\|\cdot\|_\infty$ , is a Banach space and, thus,  $S_R^n$ , with norm  $\|(\alpha^n, v^n)\| = \|\alpha^n\|_\infty + \|v^n\|_\infty$ , is a subset of a Banach space. Similarly,  $S_R$  with norm  $\sum_n \|(\alpha^n, v^n)\|$  is a subset of a Banach space.

Let us show that for each  $n$ , the (nonempty) set  $S_R^n$  is convex and compact. To prove convexity, let  $(\alpha_1^n, v_1^n)$  and  $(\alpha_2^n, v_2^n) \in S_R^n$ , let  $\delta \in [0, 1]$ , and let  $(\alpha_3^n, v_3^n) = (\delta\alpha_1^n + (1 - \delta)\alpha_2^n, \delta v_1^n + (1 - \delta)v_2^n)$ . Then  $(\alpha_3^n, v_3^n)$  satisfies the integral condition  $v_3^n(\theta) = v_3^n(0) + \int_{\theta}^1 \alpha_3^n(\tau) d\tau$ , since integration is a linear operator, so that  $(\alpha_3^n, v_3^n) \in S^n$ , and it is direct that  $(\alpha_3^n, v_3^n)$  satisfies Cn1–Cn4.

To prove compactness, let  $(\alpha^{n,k}, v^{n,k})_{k=1}^\infty$  be a sequence of elements of  $S_R^n$ . By Cn1 and the definition of  $b_1$ , we have by Cn3 that for each  $k$ ,  $0 \leq \alpha^{n,k}(0) \leq \alpha^{n,k}(1) \leq b_1$ . It follows using Cn3 and the Arzela–Ascoli theorem (for example, Rudin (1987, Theorem 11.28, p. 245)) that there exists  $\alpha^n$  satisfying Cn1 and Cn3, and a subsequence along which  $\|\alpha^{n,k} - \alpha^n\|_\infty \rightarrow 0$ . Note that  $\alpha^n$  is increasing and has range contained in  $[0, b_1]$ , and so is integrable. Since  $v^{n,k}(1)$  lies in a compact set by Cn2 and Cn4, we can, taking a further subsequence and re-indexing, assume that along the chosen subsequence  $v^{n,k}(1) \rightarrow \bar{v}$  for some  $\bar{v}$ . For each  $\theta \in [0, 1]$ , define  $v^n(\theta) = \bar{v} - \int_{\theta}^1 \alpha^n(\tau) d\tau$ . We claim that (i) along the same subsequence,  $\|v^{n,k} - v^n\|_\infty \rightarrow 0$ , and (ii)  $(\alpha^n, v^n) \in S_R^n$ . To see (i), note that for each

$\theta$  and  $k$ ,  $v^{n,k}(\theta) = v^{n,k}(1) - \int_{\theta}^1 \alpha_k^n(\tau) d\tau$ , and, hence,

$$|v^n(\theta) - v^{n,k}(\theta)| \leq |v^{n,k}(1) - \bar{v}| + \int_{\theta}^1 |\alpha^{n,k}(\tau) - \alpha^n(\tau)| d\tau \leq |v^{n,k}(1) - \bar{v}| + \|\alpha^{n,k} - \alpha^n\|_{\infty}$$

and so, since this is independent of  $\theta$ , but converges to zero,  $\|v_k^n - v^n\|_{\infty} \rightarrow 0$ . To see (ii), note that we have checked Cn1 and Cn3, and that weak inequalities are preserved under limits, and so Cn2 and Cn4 hold as well. Thus,  $S_R^n$  is sequentially compact and so, as a metric space, is compact.

Since  $N$  is finite and each  $S_R^n$  is nonempty, convex, and compact, so is  $S_R = \times_{n=1}^N S_R^n$ . Fix  $s \in S_R$ , let  $s^k \rightarrow s$ , and fix  $n$ . By stacking and since  $s \in S_R$ , there exist  $\theta_l$  and  $\theta_h$  such that  $\varphi(\theta, s) = 1$  on  $(\theta_l, \theta_h)$  and  $\varphi(\theta, s) = 0$  for  $\theta \notin [\theta_l, \theta_h]$ . But then, since for each  $n'$ ,  $\|v^{n',k} - v^{n'}\| \rightarrow 0$ , and again using stacking, for any given  $\delta > 0$  and for  $s'$  close enough to  $s$ ,  $\varphi(\theta, s') = 1$  on  $[\theta_l + \delta, \theta_h - \delta]$  and  $\varphi(\theta, s') = 0$  for  $\theta \notin (\theta_l - \delta, \theta_h + \delta)$ . Since  $\|\alpha^{n,k} - \alpha^n\| \rightarrow 0$  as well, and since  $\pi$  is bounded and continuous,  $\Pi^n(s^k) \rightarrow \Pi^n(s)$ , and, thus, that  $\Pi^n$  is continuous on  $S_R$ .<sup>52</sup>

Fix  $n$ . Since  $\Pi^n$  is continuous on  $S_R$ , and since  $S_R^n$  is nonempty, compact, and independent of  $s^{-n}$ , the theorem of the maximum implies that  $BR_R^n(s^{-n}) = \arg \max_{s^n \in S_R^n} \Pi^n(s^n, s^{-n})$  is nonempty and compact-valued for each  $s^{-n}$ , and is upper hemicontinuous in  $s^{-n}$ .

Finally, let us show that  $BR_R^n(s^{-n})$  is convex. Let  $\hat{s}^n \in BR_R^n(s^{-n})$ , with single-dominance region  $(\hat{\theta}_l, \hat{\theta}_h)$ . Then, by Corollary 1,  $(\hat{\theta}_l, \hat{\theta}_h)$  maximizes  $r$ , and on  $(\hat{\theta}_l, \hat{\theta}_h)$ ,  $\hat{s}^n = \tilde{s}(\hat{\theta}_l, \hat{\theta}_h)$ , and by Lemma 20,  $(\hat{\theta}_l, \hat{\theta}_h) = (\lambda(\theta_h^*), \theta_h^*)$ . Thus, any two elements of  $BR_R^n(s^{-n})$  win for sure on  $(\lambda(\theta_h^*), \theta_h^*)$ , agree with  $\tilde{s}(\lambda(\theta_h^*), \theta_h^*)$  on  $(\lambda(\theta_h^*), \theta_h^*)$ , and lose for sure for  $\theta \notin [\lambda(\theta_h^*), \theta_h^*]$ . But then their convex combination does the same and so is also a best response.

We have shown that  $S_R$  is a nonempty, compact, convex subset of a Banach space, and that the correspondence defined by  $BR_R(s) \equiv BR_R^1(s^{-1}) \times \dots \times BR_R^N(s^{-N})$  from  $S_R$  to  $S_R$  has a closed graph and nonempty convex values. Thus, by the Kakutani–Fan–Glicksberg theorem (Aliprantis and Border (2006, Corollary 17.55, p. 583)),  $BR_R$  has a fixed point on  $S_R$ , and we are done. Q.E.D.

#### APPENDIX D: PROOF FOR SECTION 7

We apply Reny (1999, Corollary 5.2). We first show that  $W^n$  is nonempty and compact. Let  $\hat{W}^n$  be defined as was  $W^n$  except that instead of  $v^n(\theta) \leq q^n(\theta, v^n)$ , we impose  $v^n(\theta) \leq \max_{a \in [0, \bar{a}]} (V^n(a) + \bar{a})$ . As a set of functions with uniform upper and lower bound and uniform Lipschitz bound,  $\hat{W}^n$  is compact in the uniform topology, and since  $q^n(\theta, v^n) \leq \max_{a \in [0, \bar{a}]} (V^n(a) + \bar{a})$ ,  $W^n \subseteq \hat{W}^n$ . Hence, it is enough to show that  $W^n$  is closed. But

$$G(\theta, v^n) = \{a \in [0, \bar{a}] | v^n(\theta) + a(\theta' - \theta) - v^n(\theta') \leq 0 \forall \theta' \in [0, 1]\},$$

and so  $G$  is defined by a set of weak inequalities of continuous functions and, hence, is upper hemicontinuous. But then, from Aliprantis and Border (2006, Lemma 17.30, p. 569),  $q^n$  is upper semicontinuous and so  $\hat{W}^n$  is closed.

<sup>52</sup>Recall that without stacking and outside of  $S_R$ , it is easy to construct examples where payoffs are discontinuous.



Hence,  $(W, \Pi_e)$  is a compact Hausdorff game. By **Reny (1999, Corollary 5.2)**, it is thus enough to show that  $(\bar{W}, \bar{\Pi}_e)$  is both reciprocally upper semicontinuous and payoff secure. Given efficient tie-breaking, reciprocal upper semicontinuity follows from **Reny (1999, Proposition 5.1)**. Indeed, if  $\mathcal{N}(\theta, v) = \{n | v^n(\theta) \geq v^{n'}(\theta) \forall n'\}$  is the set of firms offering maximal surplus at  $\theta$ , then by efficient tie-breaking, the sum of payoffs at  $\theta$  is  $\max_{n \in \mathcal{N}(\theta, v)} (q^n(\theta, v^n) - v^n(\theta))$ , since, among  $n \in \mathcal{N}(\theta, v)$ , the type is allocated to one for whom  $q^n$  is maximized. But then, since  $\mathcal{N}$  is upper hemicontinuous and  $q^n(\theta, v^n)$  is upper semicontinuous, the sum of the payoffs at  $\theta$  is upper semicontinuous. Integrating across  $\theta$  yields the result.

Let us turn to payoff security. The game  $(\bar{W}, \bar{\Pi}_e)$  is payoff secure if for each strategy profile  $\mu \in \bar{W}$  and each  $\varepsilon' > 0$ , each firm  $n$  has a strategy  $\hat{\mu}^n \in \bar{W}^n$  such that  $\bar{\Pi}_e^n(\hat{\mu}^n, \hat{\mu}^{-n}) \geq (1 - \varepsilon')\bar{\Pi}_e^n(\mu)$  for all  $\hat{\mu}^{-n}$  in some open ball around  $\mu^{-n}$ . Define  $\tau = \max_{n, v \in W} \Pi_e^n(v) \leq \max_{n, a \in [0, \bar{a}]} (V^n(a) + \bar{a} - \bar{u}) < \infty$ . Fix  $\mu^{-n} \in \bar{W}^{-n}$ ,  $v^n \in W^n$ , and  $\varepsilon > 0$ . Letting  $\delta_{v^n} \in \bar{W}^n$  be the Dirac measure putting probability 1 on the pure strategy  $v^n$ , we will show that by using  $\delta_{v^n+4\varepsilon}$ ,  $n$  can secure a payoff of  $\bar{\Pi}_e^n(\delta_{v^n}, \mu^{-n}) - (5 + \tau)\varepsilon$ . Integration with respect to  $\mu^n$  and  $\varepsilon = \varepsilon'/(5 + \tau)$  yields the result.

For each  $v^{-n} \in W^{-n}$ , let  $B_\varepsilon(v^{-n})$  be the open  $\varepsilon$ -ball around  $v^{-n}$ , with boundary  $\partial B_\varepsilon(v^{-n})$ . The collection of such balls is an open cover of the compact set  $W^{-n}$ , and so there is a finite set  $C \subseteq W^{-n}$  such that  $\{B_\varepsilon(c)\}_{c \in C}$  is a subcover of  $W^{-n}$ . For each  $c \in C$   $W^{-n} = \bigcup_{\varepsilon'' \geq 0} (\partial B_{\varepsilon''}(c))$ , there is at most a countable set of  $\varepsilon'' > 0$ , where  $\mu^{-n}(\partial B_{\varepsilon''}(c)) > 0$ . Hence, since  $C$  is finite, there is  $\hat{\varepsilon}$  in  $[\varepsilon, 2\varepsilon)$  such that  $\mu^{-n}(\tilde{\partial}) = 0$ , where  $\tilde{\partial} \equiv \bigcup_{c \in C} \partial B_{\hat{\varepsilon}}(c)$ .

Draw the Venn diagram on  $W^{-n} \setminus \tilde{\partial}$  corresponding to the set of balls  $B_{\hat{\varepsilon}}(c)$  for  $c \in C$ , with subsets  $W_1^{-n}, \dots, W_{M^*}^{-n}$ , where  $M^* = 2^M - 1 < \infty$ . That is, any two points in  $W^{-n} \setminus \tilde{\partial}$  are in the same  $W_i^{-n}$  if and only if the set of  $c \in C$  for which they are within  $B_{\hat{\varepsilon}}(c)$  is the same. Each  $W_i^{-n}$  is open, since  $\tilde{\partial}$  is excluded, and so each  $v^{-n}$  in  $W^{-n} \setminus \tilde{\partial}$  and  $c$  in  $C$  is either strictly less or strictly greater than  $\hat{\varepsilon}$  apart, and strict inequalities hold on a neighborhood. The sets  $W_i^{-n}$  are disjoint, with  $\bigcup_{i=1}^{M^*} W_i^{-n} = W^{-n} \setminus \tilde{\partial}$ .

For  $i \in \{1, \dots, M^*\}$ , let  $w_i = \sup_{v^{-n} \in W_i^{-n}} \Pi_e^n(v^n, v^{-n})$  bound the profit  $n$  can attain using  $v^n$ , given that  $v^{-n} \in W_i^{-n}$ . Since  $0 \leq q^n(\theta, v^n) - v^n(\theta) < \infty$  by construction of  $W^n$ ,  $0 \leq w_i < \infty$ . Let  $v_i^{-n} \in W_i^{-n}$  come within  $\varepsilon$  of attaining  $w_i$ . For any given  $\theta$ , if  $v^n(\theta) \geq v_i^{-n}(\theta)$ , then  $v^n(\theta) + 4\varepsilon > v^{-n}(\theta)$  for any  $v^{-n} \in W_i^{-n}$ , since  $W_i^{-n}$  has diameter at most  $2\hat{\varepsilon} < 4\varepsilon$ . Hence,  $\varphi_e^n(\cdot, (v^n + 4\varepsilon, v^{-n})) \geq \varphi_e^n(\cdot, (v^n, v_i^{-n}))$ . Further,  $q^n(\theta, v^n) = q^n(\theta, v^n + 4\varepsilon)$ , since  $v^n$  and  $v^n + 4\varepsilon$  have the same subdifferential. Hence, since  $q^n(\theta, v^n) - v^n(\theta) \geq 0$  and for any  $\hat{\mu}^{-n}$ ,

$$\begin{aligned} & \bar{\Pi}_e^n(\delta_{v^n+4\varepsilon}, \hat{\mu}^{-n}) \\ &= \int_{W^{-n}} \left( \int (q^n(\theta, v^n + 4\varepsilon) - v^n(\theta) - 4\varepsilon) \varphi_e^n(\theta, (v^n + 4\varepsilon, v^{-n})) h(\theta) d\theta \right) d\hat{\mu}^{-n}(v^{-n}) \\ &\geq -4\varepsilon + \int_{W^{-n} \setminus \tilde{\partial}} \left( \int (q^n(\theta, v^n) - v^n(\theta)) \varphi_e^n(\theta, (v^n + 4\varepsilon, v^{-n})) h(\theta) d\theta \right) d\hat{\mu}^{-n}(v^{-n}) \\ &\geq -4\varepsilon + \int_{W^{-n} \setminus \tilde{\partial}} \left( \int (q^n(\theta, v^n) - v^n(\theta)) \varphi_e^n(\theta, (v^n, v_i^{-n})) h(\theta) d\theta \right) d\hat{\mu}^{-n}(v^{-n}) \end{aligned}$$

$$\begin{aligned} &\geq -4\varepsilon + \sum_{i=1}^{M^*} \int_{W_i^{-n}} (w_i - \varepsilon) d\hat{\mu}^{-n}(v^{-n}) \\ &\geq -5\varepsilon + \sum_{i=1}^{M^*} w_i \hat{\mu}^{-n}(W_i^{-n}), \end{aligned}$$

where the first inequality first moves at most  $-4\varepsilon$  outside of the integral and then eliminates  $\tilde{\delta}$ , the second inequality uses that  $\varphi_e^n(\theta, (v^n + 4\varepsilon, v^{-n})) \geq \varphi_e^n(\theta, (v^n, v_i^{-n}))$ , and the third one uses the definition of  $v_i^{-n}$ .

To complete the proof, choose  $\lambda > 0$  such that for all  $\hat{\mu}^{-n} \in B_\lambda(\mu^{-n})$ ,  $\hat{\mu}^{-n}(W_i^{-n}) \geq (1 - \varepsilon)\mu^{-n}(W_i^{-n})$  for each  $1 \leq i \leq M^*$ .<sup>53</sup> Then, for all  $\hat{\mu}^{-n} \in B_\lambda(\mu^{-n})$ , and since  $w_i \geq 0$ ,

$$\begin{aligned} \bar{\Pi}_e(\delta_{v^n+4\varepsilon}, \hat{\mu}^{-n}) &\geq -5\varepsilon + \sum_{i=1}^{M^*} w_i \hat{\mu}^{-n}(W_i^{-n}) \\ &\geq -5\varepsilon + (1 - \varepsilon) \sum_{i=1}^{M^*} w_i \mu^{-n}(W_i^{-n}) \\ &\geq -5\varepsilon + (1 - \varepsilon) \bar{\Pi}_e^n(\delta_{v^n}, \mu^{-n}) \geq -\varepsilon(5 + \tau) + \bar{\Pi}_e^n(\delta_{v^n}, \mu^{-n}), \end{aligned}$$

where the third inequality uses  $\sum_{i=1}^{M^*} \mu^{-n}(W_i^{-n}) = 1$ . Q.E.D.

REFERENCES

ALIPRANTIS, C. D., AND K. C. BORDER (2006): *Infinite Dimensional Analysis*. Springer Verlag, Berlin. [1285, 1308]  
 ARMSTRONG, M., AND J. VICKERS (2001): “Competitive Price Discrimination,” *Rand Journal of Economics*, 32 (4), 579–605. [1271]  
 BAUMOL, W. (1988): *Contestable Markets and the Theory of Industry Structure*. Harcourt College Pub. [1274]  
 BENABOU, R., AND J. TIROLE (2016): “Bonus Culture: Competitive Pay, Screening, and Multitasking,” *Journal of Political Economy*, 124 (2), 305–370. [1269]  
 BIGLAISER, G., AND C. MEZZETTI (1993): “Principals Competing for an Agent in the Presence of Adverse Selection and Moral Hazard,” *Journal of Economic Theory*, 61 (2), 302–330. [1266,1269,1273,1278]  
 BILLINGSLEY, P. (2013): *Convergence of Probability Measures*. John Wiley & Sons. [1310]  
 CHADE, H., J. SWINKELS (2021): “Supplement to ‘Screening in Vertical Oligopolies,’” *Econometrica Supplemental Material*, 89, <https://doi.org/10.3982/ECTA17016>. [1273]  
 CHAMBERLAND, M. (2015): *Single Digits: in Praise of Small Numbers*. Princeton University Press. [1287]  
 CHAMPSAUR, P., AND J.-C. ROCHET (1989): “Multiproduct Duopolists,” *Econometrica*, 57 (3), 533–557. [1269]  
 CHOI, M., AND L. SMITH (2017): “Ordinal Aggregation Results via Karlin’s Variation Diminishing Property,” *Journal of Economic theory*, 168, 1–11. [1289]  
 CRAWFORD, V. P., AND E. M. KNOER (1981): “Job Matching With Heterogeneous Firms and Workers,” *Econometrica*, 49 (2), 437–450. [1270]  
 BECKHOUT, J., AND P. KIRCHER (2018): “Assortative Matching With Large Firms,” *Econometrica*, 86, 85–132. [1270]  
 EPSTEIN, L., AND M. PETERS (1999): “A Revelation Principle for Competing Mechanisms,” *Journal of Economic theory*, 88, 119–160. [1266]  
 GUERRIERI, V., R. SHIMER, AND R. WRIGHT (2010): “Adverse Selection in Competitive Search Equilibrium,” *Econometrica*, 78 (6), 1823–1862. [1270]

<sup>53</sup>If such a  $\lambda$  did not exist, one could find  $i$  and a sequence  $\hat{\mu}_k^{-n} \rightarrow \mu^{-n}$  such that  $\liminf_k \hat{\mu}_k^{-n}(W_i^{-n}) \leq (1 - \varepsilon)\mu^{-n}(W_i^{-n})$ . But, by the portmanteau theorem (Billingsley (2013, Theorem 2.1(iv))), since  $W_i^{-n}$  is open,  $\liminf_k \hat{\mu}_k^{-n}(W_i^{-n}) \geq \mu^{-n}(W_i^{-n})$ , a contradiction.

- JULLIEN, B. (2000): "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93 (1), 1–47. [1266,1269,1272-1275,1277]
- KELSO, A. S., AND V. P. CRAWFORD (1982): "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, 50 (6), 1483–1504. [1270]
- LAFFONT, J.-J., AND D. MARTIMORT (2002): *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press. [1269]
- LESTER, B., A. SHOURIDEH, V. VENKATESWARAN, AND A. ZETLIN-JONES (2018): "Screening and Adverse Selection in Frictional Markets," *Journal of Political Economy*, (forthcoming). [1270]
- LIU, Q., G. J. MAILATH, A. POSTLEWAITE, AND L. SAMUELSON (2014): "Stable Matching With Incomplete Information," *Econometrica*, 82 (2), 541–587. [1270]
- MAGGI, G., AND A. RODRIGUEZ-CLARE (1995): "On Countervailing Incentives," *Journal of Economic Theory*, 66 (1), 238–263. [1269]
- MARTIMORT, D., AND L. STOLE (2002): "The Revelation and Delegation Principles in Common Agency Games," *Econometrica*, 70, 1659–1673. [1266]
- MASKIN, E., AND J. RILEY (1984): "Monopoly With Incomplete Information," *The RAND Journal of Economics*, 15 (2), 171–196. [1265,1269,1275,1289]
- MUSSA, M., AND S. ROSEN (1978): "Monopoly and Product Quality," *Journal of Economic theory*, 18 (2), 301–317. [1265,1269,1275,1289]
- QUAH, J., AND B. STRULOVICI (2012): "Aggregating the Single Crossing Property," *Econometrica*, 80, 2333–2348. [1289]
- RENY, P. (1999): "On the Existence of Pure and Mixed Strategy Nash Equilibria in Discontinuous Games," *Econometrica*, 67, 1029–1056. [1284,1289,1308,1309]
- ROCHET, J.-C., AND L. STOLE (2002): "Nonlinear Pricing With Random Participation," *The Review of Economic Studies*, 69, 277–311. [1271]
- ROTHSCHILD, M., AND J. STIGLITZ (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 90 (4), 629–649. [1265,1270]
- RUDIN, W. (1987): *Real and Complex Analysis* (Third Ed.). McGraw Hill. [1307]
- SALOP, S. (1979): "Monopolistic Competition With Outside Goods," *The Bell Journal of Economics*, 10 (1), 141–156. [1269]
- SPULBER, D. F. (1989): "Product Variety and Competitive Discounts," *Journal of Economic Theory*, 48 (2), 510–525. [1269]
- STOLE, L. A. (1995): "Nonlinear Pricing and Oligopoly," *Journal of Economics & Management Strategy*, 4 (4), 529–562. [1269]
- (2007): "Price Discrimination and Competition," *Handbook of industrial organization*, 3, 2221–2299. [1269]
- TIROLE, J. (1988): *Industrial Organization*. MIT Press. [1267]

---

*Co-editor Alessandro Lizzeri handled this manuscript.*

*Manuscript received 25 January, 2019; final version accepted 30 October, 2020; available online 11 November, 2020.*