Screening in Vertical Oligopolies

Hector Chade*and Jeroen Swinkels†

May 2020

Abstract

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 upwards 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-stategy 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 multiplant 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.

*Arizona State University, hector.chade@asu.edu

†Northwestern University, j-swinkels@northwestern.edu

We thank the Co-Editor and 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 SED–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, and Yale University for their helpful comments.

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. 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, these choices also affect how it competes with the artisans of Hermès above them and deep supply chains of Coach below. Delta Airlines offers its customers a multitude of quality levels, but competes with discount airlines below and private jets above. Consumer-packaged-goods firms sell products at multiple quality and price points, but in a competitive environment. Consulting firms screen their workers into appropriate roles, but also compete for talent with rivals.

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

This paper helps fill this gap. An oligopoly with a finite set of vertically differentiated firms faces a continuum of agents with quasilinear preferences and private information about their ability in a labor market, or their willingness to pay for quality or quantity in a product market.[1] 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.[2] Agents then choose the firm and contract that suits them best, resolving ties across firms equiprobably. We focus on pure strategy Nash equilibria.

We first derive a set of properties that any equilibrium exhibits.[3] Our model has private values–the type of an agent enters the firm's profit only through the contract chosen. Hence, firms

[1] A model with both vertical and horizontal differentiation would also be of great interest.

[2] This is not without loss of generality (Epstein and Peters (1999), Martimort and Stole (2002)), but is economically reasonable in most settings. By Corollary 1 in Martimort and Stole (2002), there is no further loss in assuming that firms simply post menus as they do here. See also the "Extended Example" in their Section 5.

[3] As discussed below, several of these necessary conditions are closely related to ones that Jullien (2000) derives in the case of single principal who faces a type-dependent participation constraint.

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, and a more efficient 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, then adjacent firms may also tie on an interval of types at their shared boundary, and offer a zero-profit contract to those types. Positive sorting is more subtle under incomplete information, and 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 trade-off between efficiency and information rents that reflects that the firm serves only a segment of the market and in general faces a binding participation constraint both at 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 types.

These conditions yield a clear pattern of distortions. The highest firm distorts all actions *downwards*–lowering the action of a type lowers the information rents of higher types. In turn, the lowest firm distorts effort *upwards* 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 *downwards* for types below a threshold, and *upwards* 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 (1962) argues that a rail company provides roofless carriages in third class to "frighten the rich." This only reduces the amount that can be charged to the poor a little, and helps to sell second-class seats. But Dupuit also argues that first-class passengers receive "superfluous" quality. And indeed, if the firm competes against higher quality alternatives, then for high types, an inefficiently high quality can be largely reflected in the price, but the high price helps to sell second-class seats.

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. An equilibrium is thus characterized by a numerically tractable set of equations, which we exploit for examples and exploration. Second, an equilibrium in pure strategies *exists*.[4]

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 non-trivial and complicates the use of off-the-shelf existence results. The key is to reparameterize our problem 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 this type can provide. This generates intervals of types who prefer complete information. Indeed, *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 is pro-consumer, effectively restoring asymmetric information. But, if there is a capable competitor in an adjacent segment (perhaps Walmart in e-commerce) then allowing firms to tailor offers may incentivize them to compete aggressively on a broader array of customers.

We then 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 is of the order $1/N^2$, where the first N captures the extent of differentiation, while the second that when market share is small, the trade-off in chasing extra market share is favorable.

Consider next mergers. Holding fixed the behavior of other firms, and even if forced to serve the same types, a merged firm will reconfigure its action profile to reduce the information rents of its interior types. To protect customers post-merger, it is not enough to insist that the firm does not shed customers; the well-being of interior customers must also be addressed. We then show that the firm, holding fixed the behavior of other firms, will wish to shed types. The proof is non-trivial, because the firm is simultaneously adjusting its action profile. In equilibrium, non-merged

[4] Existence of a sensible pure strategy equilibrium when firms are less differentiated is an open question.

firms also adjust their behavior. It is intuitive, and true in 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. But, 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.

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 adverse selection (see Mussa and Rosen (1978) and Maskin and Riley (1984), and 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. 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. Matching is at the heart of our analysis. Biglaiser and Mezzetti (1993) analyze a versioin of our setup with two firms. Ties are broken in favor of a firm that gains the most from that type. This tames payoff discontinuities, but is less economically natural. Much of economic interest requires that we move beyond two firms.

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, and 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 cannot 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 pattern of 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 option of the agent. As a maximum of two functions, the new technology fails concavity. Our results allow partial participation.[5]

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).[6] 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 differ in a parameter $\theta \in [0,1]$ with cumulative distribution function (cdf) H with strictly positive and $\mathcal{C}^1$ density h.[7] We assume H and $1-H$ are strictly log-concave.[8]

The agent chooses an action $a \geq 0$. The value before transfers of action a is $\mathcal{V}^n(a)$ for firm n, and $\mathcal{V}(a,\theta)$ for the agent of type θ. 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 $\mathcal{V}(a,\theta) = \hat{\mathcal{V}}(a) + a\theta$. 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 $\mathcal{V}^n(a) + t$, $\mathcal{V}(a,\theta) - t$, where t would typically be positive in a product market, where the agent is a customer, and negative in a labor market, where the agent is a worker. For simplicity, for much of the paper we assume that the agent has no outside option beyond the offers of the various firms.[9] 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 $V^n(a) = \mathcal{V}^n(a) + \hat{\mathcal{V}}(a)$, so that $V^n(a) + a\theta$ is the match surplus between Firm n and type θ when action a is taken. For each n, V^n is strictly concave, with $V^n_{aa} < 0$, $V^n_{zz} < 0$, and the determinant of the Hessian of V^n strictly positive. We also assume that $V^n_a(0) \geq 1/h(0)$ and $\lim_{a\to\infty} V^n_a(a) < -1 - (1/h(1))$, which will ensure that actions are interior.

Example 1 ***A product market with quality differentiation.*** *Let* $\mathcal{V}(a,\theta) = \sqrt{\rho + a} + a\theta$*,*

[5] 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.

[6] Another example with sorting and incomplete information is Liu, Mailath, Postlewaite, and Samuelson (2014).

[7] We use increasing and decreasing in the weak sense, adding 'strictly' when needed, and similarly with positive and negative, and concave and convex. We write $(f)_x$ for the total derivative of f with respect to x. We use the symbol $=_s$ 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.

[8] Our model is equivalent to one with a single agent drawn from H. Log-concavity is standard, and avoids the need for ironing techniques.

[9] It is easy to include a type-independent outside option. On a type-dependent outside option, see Section 5.3.

where $\rho > 0$ is sufficiently small, be the value to the customer of type θ 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 convex, strictly submodular in n and a, and satisfies $\lim_{a\to\infty} c_a^n(a) = \infty$. Here, higher indexed firms have lower marginal costs for producing quality.

Example 2 ***A product market with quantity differentiation.*** *Let h be uniform, and let the customer have demand curve $P(a) = 3 + \theta - a$, so that $\mathcal{V}(a, \theta) = (3 + \theta)a - (a^2/2)$. Let the firm have fixed cost ζ^n per customer, and variable cost β^n per unit sold, where $\beta^n \leq 2$ is strictly decreasing in n, so that $\mathcal{V}^n(a) = -\zeta^n - \beta^n a$, and thus higher indexed firms have lower marginal cost of the quantity consumed by any given customer.*

Example 3 ***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 $\mathcal{V}_a^n(0) \geq 1/h(0)$ and $\lim_{a\to\infty} \mathcal{V}^n(a) = 0$. For example, let $\mathcal{V}^n(a) = \zeta^n + \beta^n \log(\rho + a)$, where $\rho > 0$ is sufficiently small, and β^n is strictly increasing in n, so that higher indexed firms value effort more. Let the cost to the worker be $c(a) - a\theta$, where c is convex with $\lim_{a\to\infty} c_a(a) > 1 + (1/h(1))$, so that $\mathcal{V}(a, \theta) = -c(a) + a\theta$.*

Let $v_*^n(\theta) = \max_a V^n(a) + a\theta$ be the most surplus Firm n can offer type θ without losing money, and let $\alpha_*^n(\theta)$ be the associated maximizer. We assume that each firm n is *relevant* in that there is θ such that $v_*^n(\theta) > \max_{n'\neq n} v_*^{n'}(\theta)$. We will see that relevance is sufficient for all firms to be active in equilibrium. For the two parameterized examples above, this condition is satisfied by appropriate choice of ζ^n.[10] By relevance, and using strict supermodularity of $V^n(a)$, for each firm n, there is an open interval (a_e^{n-1}, a_e^n) of actions such that n is the most efficient firm at action a, where for $1 \leq n < N$, $V^n(a_e^n) = V^{n+1}(a_e^n)$, with $a_e^0 = 0$ and $a_e^N = \infty$.[11]

Firms simultaneously offer menus of contracts, where Firm n's menu is a pair of functions (α^n, t^n), with $\alpha^n(\theta)$ the action required of an agent who chooses Firm n and announces type θ, and $t^n(\theta)$ the transfer to the 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

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

It is without loss that firms offer incentive compatible menus. Thus, a menu can equally well be described as (α^n, v^n), where, as is standard, incentive compatibility is equivalent to requiring that the action schedule α^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

[10] See Section 5.2 for a more general way to generate economies where all firms are relevant.

[11] 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 non-empty 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 below that (a_e^{n-1}, a_e^n) non-empty is necessary.

all θ (so in particular, v^n is convex). We will do so henceforth.[12] Since α^n is increasing, it jumps at at most a countable set of points, and so since h is atomless, it is without loss to assume α^n at any $\theta < 1$ to be right continuous and α^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 $s^{-n} \in \times_{n' \neq n} S^{n'}$ be a typical strategy profile for firms other than n. Firm n's profit on a type-θ 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 (α, v), we write $\pi^n(\theta, \alpha, v)$ for $\pi^n(\theta, \alpha(\theta), v(\theta))$.

After observing the posted menus, the agents sort themselves to the most advantageous firm. Formally, for any n and $s^{-n} \in S^{-n}$, 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_θ^{-n}. From the point of view of n, (a^{-n}, v^{-n}) summarizes all the relevant information about the strategy profile of his opponents. Define $\varphi^n(\theta, s)$ as the probability that n serves θ 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.

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 φ, we can view the game as simply one among the firms, with strategy set S^n and payoff function Π^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 Nash equilibrium (in pure strategies) 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 characterize equilibrium. We begin with a set of necessary conditions. Then, we show that a subset of these conditions is sufficient under a restriction, *stacking*, that the firms are sufficiently differentiated. Leveraging this, we state an existence result. Finally, we show how to use our conditions to analyze the model numerically, a surprisingly tractable exercise.

4.1 Necessity

We begin with a set of necessary conditions for a Nash equilibrium in pure strategies. We first state the main theorem. Then, we define each of the terms, and, for each necessary condition, argue at an intuitive level why it must hold, and flesh out some of its economic implications.

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*

[12] See Armstrong and Vickers (2001) and Rochet and Stole (2002) for other examples of competition in utility.

boundaries.

As discussed, some of the "non-equilibrium" results in this section are related to Jullien (2000), but technical differences make it hard for us to import those results off-the-shelf. We provide intuition for those results here, and full and self-contained proofs in the online appendix.

4.1.1 Positive Profits (*PP*)

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 action-surplus offerings 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.[13]

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 non-empty, a firm that offers the menu $(\alpha_*^n, v_*^n - \varepsilon)$ for ε sufficiently small will win a positive measure of agents, and earn strictly positive profits on any agents served. See Proposition 2 in Appendix A.

This result uses in an essential way that any firm that matches the most favorable offer wins with strictly positive probability. If instead, as in Biglaiser and Mezzetti (1993), 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 θ without losing money.

4.1.2 No Poaching (*NP*)

Fix an equilibrium, and let $v^O(\cdot) = \max_n v^n(\cdot)$ 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 θ,

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

so that θ 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

[13]See Jullien (2000), Lemma 3, or Online Appendix, Proposition 5.

that a firm n' could profitably out-compete n on type θ with another action, but does not do so, because it would attract some of its own types in a detrimental way.[14]

Moving v^O across the inequality, *NP* says that the second most efficient firm would lose money by poaching θ at the current action. That is, if n is winning at θ, 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 $[a_e^{n-1}, a_e^n]$.

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.[15]

4.1.3 Positive Sorting (*PS*)

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 non-empty interval (θ_l^n, θ_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 one). 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, θ' is taking an action at least as high as θ in equilibrium. But, V^n is strictly supermodular in n and a. Hence, if n sometimes serves θ' and $n' > n$ sometimes serves θ, then, by *PP* and *NP*, either n will want to always serve θ, or n' will want to always serve θ', a contradiction. The only exception is if both firms are indifferent about hiring both θ and θ', 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 (θ_l^n, θ_h^n), firm n serves each type with probability one. 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 regions 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*

[14] For *PP* and *NP* the details of tie-breaking are inessential as long as φ^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.

[15] 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.

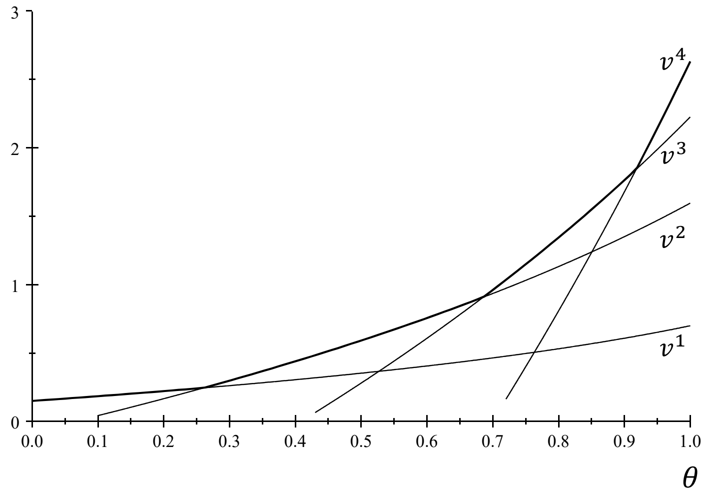

Figure 1: **An Equilibrium with Four Firms.**The curves v^n are the equilbrium surplus functions. Each firm serves the types where its surplus function is highest. The agent receives surplus indicated by the thick locus. Upward kinks in this locus reflect jumps in the action.

and four firms. We explain how the figure was generated in Section 4.3.

The conditions differ when a firm outside of the relevant interval 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 setting (see, for example, Baumol (1988)), a firm may be inactive in equilibrium, but limit active firms to offerings on which the inactive firm cannot strictly profit. Here, such a thing could occur on a type-by-type basis.[16]

4.1.4 No Extraneous Offers (*NEO*)

Let us focus on settings where *PS* holds. By *NP*, we know that with probability one, 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 α^n is continuous everywhere, and in which actions are always in $[a_e^{n-1}, a_e^n]$.[17] Appendix A shows that under *NEO*, any equilibrium has *PS*.

[16]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.

[17]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.

4.1.5 Internal Optimality (*IO*)

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 γ^n by

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

Strategy profile s satisfies *internally 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 $\alpha^n(\cdot) = \gamma^n(\cdot, \kappa)$ on $[\theta_l^n, \theta_h^n]$.[18,19] By *IO*, there is a type $\theta_0^n \in [\theta_l^n, \theta_h^n]$ satisfying $H(\theta_0^n) = \kappa$. Actions are distorted down/up below/above θ_0^n. Actions at θ_0^n are efficient. Since $\gamma(\cdot, \kappa)$ is strictly increasing, an economic implication of *IO* is that there is complete sorting within the interval of types uniquely served by each firm.

We will shortly relate κ to a Lagrange multiplier in a suitable problem. But, to see intuition for *IO*, note that since $\kappa^N = 1$, (1) reduces 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, note that $\kappa^1 = 0$ and so π_a^1 is everywhere negative. For intermediate firms, the action is distorted down below θ_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 γ, and build structure that we will need when we turn to sufficiency and existence, fix boundary points θ_l and θ_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):

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

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

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

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

This relaxes n's problem, as we drop monotonicity of α, and ignore the outside option except at θ_l and θ_h, and relax the constraint at θ_l and θ_h. Let $\iota(\theta_l, \theta_h, \kappa) = v^{-n}(\theta_h) - v^{-n}(\theta_l) - \int_{\theta_l}^{\theta_h} \gamma(\theta, \kappa) d\theta$

[18] Since $V_a^n(0) \geq 1/h(0)$ and $\lim_{a \to \infty} V_a^n(a) < -1 - (1/h(1))$, (1) has an interior solution. By Lemma 5 in Appendix A, log-concavity of H and $1 - H$ imply that $(\kappa - H(\cdot))/h(\cdot)$ is decreasing for all $\kappa \in [0,1]$, and so, since $\pi_{a\theta} = 1 > 0$, γ^n is strictly increasing in θ. Similarly, γ^n is strictly decreasing in κ.

[19] See Jullien (2000), Theorem 1 and Proposition 2 for a similar result, and the online appendix for a self-contained proof that deals with the details of our environment.

and $\tilde{\kappa}(\theta_l,\theta_h) = \arg\min_{\kappa\in[H(\theta_l),H(\theta_h)]} |\iota(\theta_l,\theta_h,\kappa)|$.[20] 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 (θ_l,θ_h), $\tilde{\alpha}$ is uniquely defined and equal to $\gamma(\cdot,\tilde{\kappa}(\theta_l,\theta_h))$.[21] 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 η be the shadow value of increasing the surplus of type θ_h holding fixed the surplus of type θ_l. One way n can achieve this increase is to raise the action at any given interior type θ. This has benefit $\pi_a(\theta,\tilde{\alpha},\tilde{v})$ on the $h(\theta)$ types near θ, and raises the surplus on the $H(\theta_h)-H(\theta)$ types between θ and θ_h, and so $\eta = -\pi_a(\theta,\tilde{\alpha},\tilde{v})h(\theta) + H(\theta_h) - H(\theta)$, where we note that at an optimum, this expression must holds for all types, since otherwise the firm could profitably raise the action at one θ and lower it at another, leaving $v(\theta_h)$ unaffected. In particular, since $\pi_a = 0$ at θ_0, and since $H(\theta_0) = \kappa$, we have $\eta = H(\theta_h) - \kappa$. Substituting and rearranging yields (1). To use this result to show *IO*, we show that if $\alpha \neq \gamma(\cdot,\tilde{\kappa})$ then we can perturb α in the "direction" of $\gamma(\cdot,\tilde{\kappa})$ strictly profitably.[22]

4.1.6 Optimal Boundaries (*OB*)

Strategy profile s satisfies the *optimal boundary condition (OB)* if for each of $\theta = \theta_l^n$ and θ_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 $\theta_h^N = 1$.[23]

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 θ_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, θ_h increases at rate $1/(a^{-n}(\theta_h)-\alpha(\theta_h))$ since $\alpha(\theta_h)$ is the slope of v at θ_h, and $a^{-n}(\theta_h)$ is the slope of v^{-n}. Hence, the profits from the new types served is $\pi(\theta_h,\alpha,v)h(\theta_h)/(a^{-n}(\theta_h)-\alpha(\theta_h))$. Cancelling $h(\theta_h)$ and rearranging yields (5), and similarly for θ_l.

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

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

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

[20] Noting that by (1), $\gamma_\kappa < 0$, and so $\iota_\kappa > 0$, then $\tilde{\kappa}$ 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.

[21] 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 $\theta \leq \theta_l$.

[22] See the online appendix for a proof of the lemma and a proof (Proposition 6) that this implies *IO*.

[23] A close relative is Jullien (2000), Theorem 2. A proof that takes care of transversality is in the Online Appendix.

If $\alpha = \gamma(\cdot, H(\theta_0))$, then $\pi(\cdot, \alpha, v)$ is strictly single-peaked with peak at θ_0.

To see the proof of (6) note that by definition of π,

$$(\pi(\theta, \alpha, v))_\theta = \pi_\theta(\theta, \alpha, v) + \pi_a(\theta, \alpha, v)\alpha_\theta(\theta) - v_\theta(\theta),$$

and that $\pi_\theta(\theta, \alpha, v) = \alpha(\theta) = v_\theta(\theta)$. If $\alpha = \gamma(\cdot, H(\theta_0))$, then from (1), $\alpha_\theta > 0$, and $\pi_a(\theta, \alpha, v)$ has strictly the same sign as $\theta_0 - \theta$. Hence, π is strictly single-peaked at θ_0.

That profits are strictly single-peaked at θ_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 to extract rents from. Similarly, for the end firms, it is the extreme types from whom it is easiest to extract rents.

Let us now show that κ is interior for $n \notin \{1, N\}$. If $\kappa = H(\theta_h)$, then by Lemma 2, $\pi(\cdot, \alpha, v)$ is strictly increasing on (θ_l, θ_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 θ_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, π is strictly positive on (θ_l, θ_h). This follows since by *PP*, π is positive at θ_l and θ_h, and since α is of the γ form on $[\theta_l, \theta_h]$, and so π is strictly single-peaked on $[\theta_l, \theta_h]$. How about at the boundary types θ_l and θ_h? If there is a region of overlap between the two firms, then 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) will be strictly negative. Thus, $\pi(\theta_h, \alpha, v)$ must be strictly positive, and similarly for $\pi(\theta_l, \alpha, v)$. Even though firms compete for the boundary customer, the difference in their technologies implies that neither firm can profitably imitate the other.

4.2 Sufficiency and Existence

We begin by eliminating 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. The example below shows that stacking holds if firms are sufficiently differentiated.[24] 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.

[24]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 $\theta_l^2 = \theta_h^1$, and so $\alpha^1(\theta_h^1) = \gamma^1(\theta_h^1, 0) > \gamma^2(\theta_h^1, 1) = \alpha^2(\theta_h^1)$, contradicting *PS*.

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 non-local 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 faciliates numeric analysis. By *PS,* there are $N-1$ boundary points θ^n between Firms n and $n+1$, and using *IO*, each firm's behavior is characterized by a "slope" κ^n and an intercept $v^n(0)$. Since κ^1 and κ^N are fixed, we have $3N-3$ unknowns. But, at each 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 $\mathcal{V}(a, \theta) = -(3-\theta)a$. See Online Appendix, Section 2 for 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 it can undo any inefficiency, and then extract all the surplus, leaving all types worse off. In oligopoly, there is another effect: competition will increase the agents' *outside option.* With asymmetric 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 surplus equals what the second most efficient firm can provide.[25]

When comparing pure-strategy equilibria under complete and under incomplete information, a simple structure arises. For $1 \leq n < N$, let θ_*^n be the boundary point between where n and $n+1$ are the most efficient firm to serve θ. 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 θ^n be the boundary point between n and $n+1$ under incomplete information.

Theorem 3 (Welfare) *Assume stacking. Then,*

(1) For each $1 \leq n \leq N-1$, an interval of types containing θ_*^n and θ^n is strictly better off under complete information. Near θ_*^n the firm is strictly worse off under complete information;

[25]To avoid an uninteresting openness issue, we assume here that agents break ties in favor of the firm that earns more profit in serving them, and that no firm makes an offer that they would lose money on if accepted.

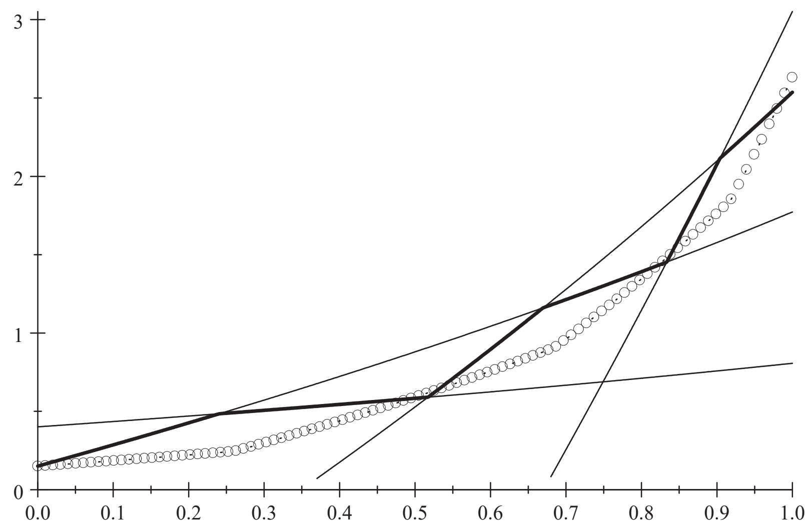

Figure 2: **Complete Information versus Incomplete Information.** The thin lines are the efficient surplus each firm can provide, v_*^n, in the setting of Figure 1. The thick line is the amount that the agent receives under complete information. The bubbled line is the surplus in the incomplete information oligopoly. Complete information is prefered 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).

(2) There is a single (possibly empty) subset of $[\theta_*^{n-1}, \theta_*^n]$ that strictly prefers being served by Firm n under incomplete information to complete information; and

(3) All types may strictly prefer complete information.

For intuition, with complete information, θ_*^n earns the efficient surplus, and the firms earn zero, and so is better off than with incomplete information. *Contrary* to monopoly, some types–and perhaps *all*–are benefited by complete information. See Figure 2, which builds on Figure 1.[26] Such examples exist for any N, although we will show next that in both cases, surplus converges to the efficient level as N grows.

5.2 The Competitive Limit

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*. In this section, we explore the behavior of these three forces as the number of firms grows.

We will consider a setting where firms can enter at a fixed cost $F > 0$, and then choose freely from a set of potential technologies parameterized by $z \in [0, \bar{z}]$. Let $V(a, z) + a\theta$ be the

[26] See the online appendix for an example where h is changed so that all types 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.

match surplus from technology z, action a, and type θ, where V is $\mathcal{C}^2$, 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 as before, that $V_a(0,0) > 1/h(0)$ and $\lim_{a\to\infty} V_a(a,\bar{z}) < -1-(1/h(1))$. In Appendix 4.1, Step 0, we show that there is a non-empty interval $[z_l, z_h]$ such that for any finite N, and for $z_l \leq z^1 < \cdots < z^N \leq z_h$, we can take $V^n(a) = V(a, z^n)$, and have a market satisfying our conditions, where z_l and z_h are the firms that are best suited to serve types 0 and 1, respectively.[27] For example, let $V(a,z) = a - a^2/2 - (a-z)^2$, take $[z_l, z_h] = [1,2]$, and take $V^n(a) = V(a, 1+n/N)$.

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 a type θ could possibly hope for is $v_*(\theta) = \max_{a,z}(V(a,z)+\theta a)$. We have the following theorem.

Theorem 4 (Limit Efficiency) *In any equilibrium with endogenous entry and NEO, there is* $\rho \in (0,\infty)$ *such that* $1/(\rho F^{1/3}) \leq N \leq (\rho/F^{1/3})+2$, *while the profit per type,* π, *and the difference between what each type* θ *earns and* $v_*(\theta)$ *are each of order* $1/N^2$.

So, as F goes to zero, the market is covered by a number of firms which grows like $1/F^{1/3}$, actions are efficient, and all surplus goes to the agents. Industry profits converge to zero like $1/N^2$, as does the total expenditure on entry costs, $N \times F$. Intuitively, gaps in the market, as measured by z, have length like $1/N$, and so, since the action is efficient for some interior type of the firm, inefficiency is of the order $1/N^2$. And, for each type, there is another firm who is nearly as efficient at the induced action, and so the gains in the market go to the agent. The proof is an exercise in Taylor expansions. At its core is that the profits of an entrant are bounded *below* by a constant times the length of the largest gap in the interval $[z_l, z_h]$ raised to the third power, and the profits of an incumbent are bounded *above* in a similar fashion, which yields the claimed relationship between N and F. See Online Appendix 4.1 for details.

5.3 Multi-Firm Monopolies and Mergers

Assume that a single firm M controls more than one V^n. For example, LVMH, through a sequence of mergers and acquisitions, controls a set of firms 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.

5.3.1 Monopoly with an Outside Option

Let M control technologies $n_l, \ldots, n_h$. Assume first that M is a monopolist facing a convex outside option $\bar{u}$ which, per Section 4.2, is first below $\gamma^{n_l}(\cdot, 1)$, and then above $\gamma^{n_h}(\cdot, 0)$.

[27] In particular, take $(z_l, z_h) = (z^\theta(0), z^\theta(1))$ as defined in Step 0.

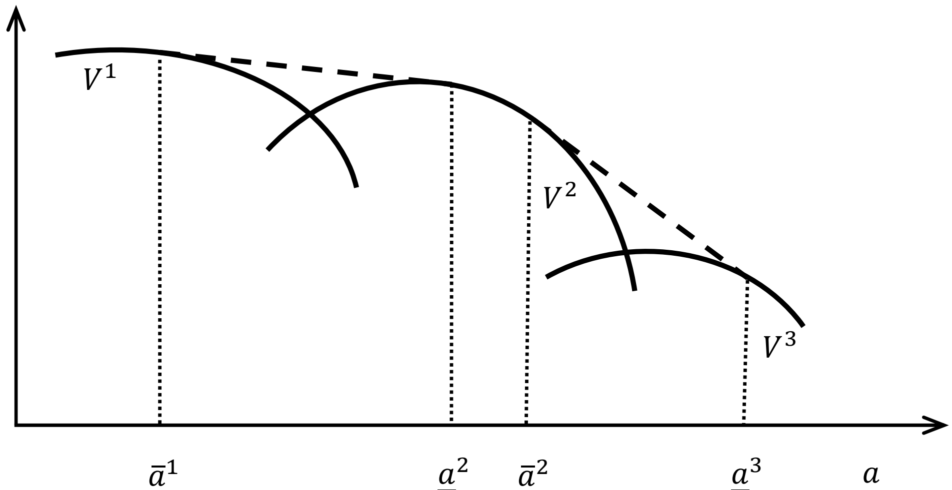

Figure 3: **Multiplant Firm.** The firm controls three technologies, V^1, V^2, and V^3. Below $\bar{a}^1$ $\bar{V} = V^1$, on $(\underline{a}^2, \bar{a}^2)$ we have $\bar{V} = V^2$, and above $\underline{a}^3$ we have $\bar{V} = V^3$. The dotted lines complete the concave envelope of the technologies.

To analyze this problem, let $\bar{V}$ be the concave envelope of $\max\{V^{n_l}, \ldots, V^{n_h}\}$ (see Figure 3). By relevance, for each $n \in \{n_l, \ldots, n_h\}$, $\bar{V}$ equals V^n over some strictly positive-lengthed interval $[\underline{a}^n, \bar{a}^n]$, where these intervals are disjoint, and where $\underline{a}^{n_l} = 0$, 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,

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

noting that $\bar{V}_a(a) + \theta$ plays the role of π_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. Again, $\gamma^M(\cdot, \kappa) = \gamma^n(\cdot, \kappa)$ and $\bar{V} = V^n$.[28]

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 type served is the same as in our previous analysis.[29] To see that this solution is also

[28]The choice of action at this finite set of points is irrelevant to the surplus integrals.

[29]Note that Firm M, in the exercise of its market power, may to idle one or more of its technologies at the top or bottom. From the construction of γ^M, a non-empty subset of consecutive technologies will be active.

optimal for M (which has technology $\max_{n\in\{n_l,\ldots,n_h\}} V^n$ rather than $\bar{V}$), note that $\bar{V}$ is at least as big as $\max_{n\in\{n_l,\ldots,n_h\}} V^n$ for all a, but the two are equal everywhere in the range of γ^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 θ and κ, 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 $\bar{u}$, and so, if n and $n+1$ remain active, then the boundary point between them, $\theta^{M,n}$, depends only on κ. Also from (7), for each θ, $\gamma^M(\theta,\kappa)$ 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

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

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

5.3.2 Oligopoly versus Monopoly with Fixed Market Size

Next, let us compare the outcome of the monopoly with a setting where firms $n_l, \ldots, n_h$ compete given $\bar{u}$.[30] In this subsection, we assume that M is forced to serve the same set of types as did the constituent firms pre-merger, but can adjust each type's action, and the allocation of types across its constituent firms. Since 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. 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 M's constituent firms. If forced to serve exactly $[\theta_l, \theta_h]$, then M will choose κ in $(\kappa^{n_l}, \kappa^{n_h})$. All types in (θ_l, θ_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 on then upwards. Firm M distorts first further downwards on a longer interval of low types, and then further upwards on a longer interval of high types. Hence, the surplus function offered by M is first flatter than in the oligopoly, and

[30] This analysis thus covers the case of an exogenous outside option $\bar{u}$, as long as that outside option is "shallow-steep," in which case the extreme firms might optimally exclude some types, but will serve an interval. With a more general outside option, a single firm might choose to serve several intervals of type, complicating the analysis.

then steeper, and so lies everywhere below it, since the two are by fiat equal at θ_l and θ_h. The proof takes into account that M will reallocate types across its constituent parts.

5.3.3 Oligopoly versus Monopoly with Endogenous Market Size

Unless legally constrained to do so, M is unlikely to serve all of $[\theta_l, \theta_h]$. Each constituent firm, knowing that it would lose types at each end, was indifferent about decreasing the surplus offered to all of its types by a small constant. But then, the merged firm–which no longer suffers the loss of types at interior boundaries–strictly prefers to lower surplus. By 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 (θ_l, θ_h) are strictly worse off compared to when the set of types served is fixed, and so, a fortiori, are strictly worse off compared to oligopoly.*

5.3.4 Oligopoly with Merged Firms

Next, 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. It is immediate that if a single firm controls two *non-sequential* sets of firms, then, in equilibrium, the competing firms in the middle will be active. Hence, *IO* and *OB* can be thought of separately for each connected set of firms, and we can without loss of generality let firm M control a sequential set of firms $n_l, \ldots, n_h$.

First fix the behavior of firms outside of $\{n_l, \ldots, n_h\}$. Then, $\bar{u}$ will 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 chooses 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 with the merged firm losing share. But, there are conflicting *economic* forces at play: when M maintains the set of types served, then from above, M moves the actions of its top an bottom agents *closer* to its nearest competitors. Hence, when those competitors raise the surplus they are offering, they gains types faster, which pushes them to fight harder for types than before, and *raise* 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. Based on numerical exploration, our speculation is that the second force dominates except perhaps in extreme examples.

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.[31] 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

[31] For numerical analysis of the merger one simply replaces the relevant instances of *OB* by (8).

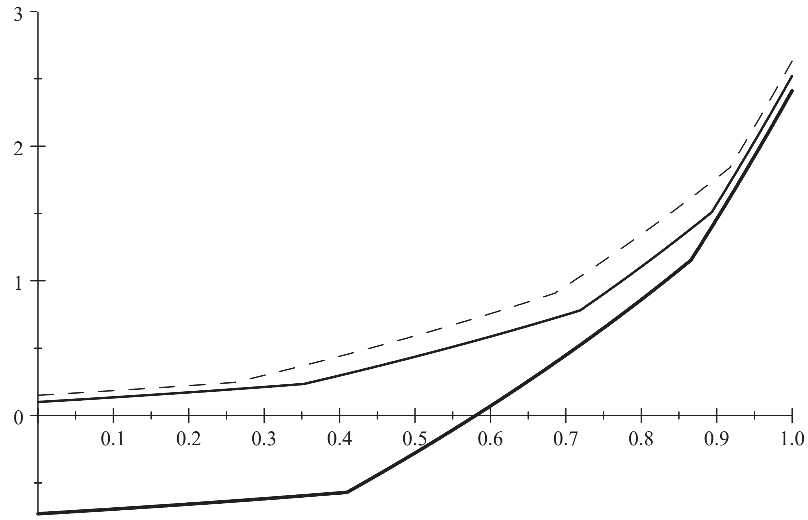

Figure 4: **A Merger and a Failing Firm.** The dashed locus is the equilibrium surplus from Figure 1. The medium weight locus is the equilibrium surplus when we put Firms 2 and 3 under the control of a single firm, M, and the heavy locus is the surplus when Firm 2 exits the market. Consumers, especially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

be absorbed by Firm 3 than to lose it altogether.[32] The merged firm competes vigorously at both ends, while when Firm 2 disappears, Firm 1 becomes more differentiated than before. 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 $\hat{s}^n$ will typically win a set of types different from either of them. But then, the first-order conditions need not imply optimality, complicating sufficiency. Existence is non-trivial because Π^n is not continuous at ties. And, since $\Pi^n(\cdot, s^{-n})$ is not quasi-concave, the set of best-responses may be non-convex, 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, which by IO will use action profiles of the γ form, concentrates simply on the choice of θ_l^n and θ_h^n. Later subsections analyze this problem and then use that analysis to prove sufficiency and existence.

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

C1 α^n is continuous, with $\alpha^n(\theta) \in [\gamma^n(\theta, 1), \gamma^n(\theta, 0)]$ for all θ.[33]

[32] This assumes that, as in this example, the merged firm chooses to operate the technology of Firm 2.
[33] We cannot just impose that firms use γ strategies, as that space is not a convex subset of the strategy space.

C2 $v^n \leq v^n_*$.

Fix n and s^{-n} satisfying $C1$ and $C2$. By relevance, Firm n earns strictly positive profits in any best response to s^{-n}. By stacking, there is $\theta^x \in [0,1]$ such that $a^{-n} < \gamma^n(\cdot, 1)$ for $\theta < \theta^x$, and $a^{-n} > \gamma^n(\cdot, 0)$ for $\theta > \theta^x$. In Figure 1, and for Firm 2, θ^x is the point at which v^1 and v^3 cross.

One of the most important implications of stacking is that *NP* dropped from the analysis, leaving only local conditions. Say that strategy s^n is *single-dominant* on (τ_l, τ_h) if $v^n > v^{-n}$ on (τ_l, τ_h), and $v^n < v^{-n}$ outside of $[\tau_l, \tau_h]$.

Lemma 3 (*OB* implies *NP*) *Assume stacking, let s satisfy* $C1$*, and assume that n sometimes wins. Then,* s^n *is single-dominant on a non-empty interval including* θ^x*, and if* s^n *satisfies OB, then it satisfies NP.*

That s^n is single-dominant on some non-empty interval including θ^x follows since by $C1$ and stacking, v^n can only cross v^{-n} once below θ^x and once above, and these crossings are strict. That *NP* is redundant follows since by $C1$, a^{-n} is above the efficient level for n to the right of θ_h, and hence by (6), the profit to poaching is decreasing. And, we show that near θ_h, *OB* implies that n does not want to poach. 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 next relate the maximization of r and Π^n, the profit in the original problem.

Proposition 1 (Equivalence) *Assume stacking. Fix n and* s^{-n} *satisfying* $C1$ *and* $C2$*. Then, r has a maximum* (θ_l, θ_h)*, and* $\hat{s}$ *is a maximum of* $\Pi^n(\cdot, s^{-n})$ *if and only if for some maximum* (θ_l, θ_h) *of r,* $\hat{s}$ *is single-dominant on* (θ_l, θ_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 relies on two lemmas in Appendix C. By Lemma 8, r has a maximum at some point (θ_l, θ_h), and the associated solution to the relaxed problem is feasible in the original game, serves the interval (θ_l, θ_h), and has the same payoff as r. By Lemma 11, for any strategy in the original game, there is (θ_l, θ_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 $C1$ and $C2$, and that any critical point of r is that maximum. We first develop some notation. Then we provide intuition and dive into the details.

We begin by showing (Lemma 12) that any optimum of r is in the rectangle $R = [0, \theta^x] \times [\theta^x, 1]$ shown in Figure 5.[34] Let K be the set of points θ where n transitions from one opponent to the

[34]For Firm 1, $\theta^x = 0$, and so the "rectangle" R becomes a vertical line segment, and similarly, for Firm N, $\theta^x = 1$, and so R becomes a horizontal line segment.

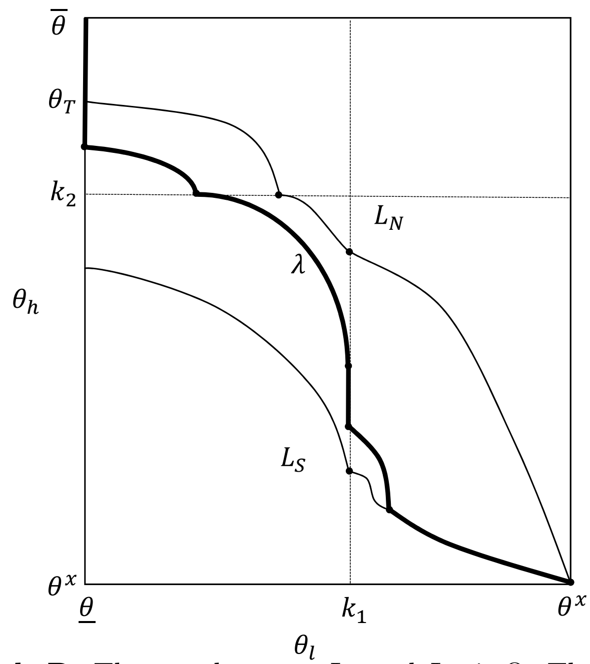

Figure 5: **The rectangle R**. The area between L_S and L_N is Θ. There are kink points in v^{-n} at k_1, k_2, and θ^x. On the four areas delineated by the dotted lines, r is continuously differentiable. The thick line is the path described by λ. Where the path runs along L_S, we have $r_{\theta_l} \leq 0$ and $r_{\theta_h} > 0$, and so ψ is increasing. The path never runs along L_N, where $r_{\theta_l} < 0$.

next (note that $|K| \leq N-2$). In Figure 5, $K = \{k_1, \theta^x, k_2\}$, so that n has two competitors below it and two above. By $C1$, each point of discontinuity of α^{-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, v^{-n} is continuously differentiable on $\tilde{R}$, and thus so is r.

6.1.1 Hiking towards 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, noting that in Figure 5, θ_l is a choice from west to east, while θ_h is a choice from south to north. This landscape has a complicated shape, with kinks and local minima. But, when the other firms play strategies satisfying $C1$ and $C2$, the firm has available positive profit strategies. So, consider the "islands" where the payoff is positive.

We show first that on each island, any place where r is differentiable in one or both directions and the first-order conditions are satisfied, is also a local maximum in those directions, so that any local minima where r is differentiable are underwater. Now, fix any latitude (a choice of θ_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 θ_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 as one move west to east is an interval. But then, there is a single island, and there is a unique path running from the south (east) to the north (west) of the island with the property that each point along the path is the highest point at that latitude. Finally, we show that payoffs are strictly quasi-concave as one hikes northward along this ridge. It follows that the island has a unique peak, and that any point that satisfies the first-order conditions is in fact that maximum.[35]

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 $C1$ 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 by $\iota(\theta_l, \theta_h, H(\theta_h)) = 0$. These are the north and south boundaries of $\Theta = \{(\theta_l, \theta_h) \in R | \iota(\theta_l, \theta_h, \tilde{\kappa}(\theta_l, \theta_h)) = 0\}$. The set Θ will be central to our analysis, because we will see shortly that any maximum of r occurs either in Θ or along a (specific) part of the boundary of R.

Section 11.2.1 begins by deriving the local properties of r. After some brush clearing, Lemma 14 shows that on any given $\tilde{R} \cap \Theta$, if $r_{\theta_l} = 0$ then r is locally strictly concave in θ_l. Similarly, if $r_{\theta_h} = 0$, then r is locally strictly concave in θ_h, and anywhere that $r_{\theta_l} = r_{\theta_h} = 0$, r is locally strictly concave in (θ_l, θ_h). Section 11.2.2 uses the local properties of r to analyze its maxima. Lemma 15 shows 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 (Assumption 1) that, as in Figure 5, 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 θ_T. Then, (Corollary 3) any maximum of r occurs either in Θ, 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 θ_h, let $\Theta(\theta_h)$ be the interval of θ_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-west. Let D be the set of θ_h such that $\psi > 0$. Fix $\theta_h \in D$ with $\theta_h < \theta_T$. Lemma 16 shows that $r(\cdot, \theta_h)$ has a unique maximum $\lambda(\theta_h)$. The proof rests on Lemma 14, but accounts for the kinks in our terrain. The path along the ridge is $(\lambda(\cdot), \cdot)$. By Lemma 16, λ is continuous, and hence so is ψ. By Lemma 17, D is an interval. The path λ never runs along L_N, because on L_N, profits strictly decrease in θ_l. Where it runs along L_S, Lemma 18 shows that ψ strictly increases.

So, consider any $\hat{\theta}_h$ such that $\lambda(\hat{\theta}_h)$ is in the interior of $\Theta(\hat{\theta}_h)$. Lemma 19 shows that the left and right derivatives of ψ at $\hat{\theta}_h$ and the left and right partial derivatives of r with respect to θ_h at

[35] Why we did not simply establish that the relevant function is strictly concave at any critical point? First, our function can have local minima "under water." Second, in $\mathbb{R}^2$ this is not enough to establish uniqueness. For example (Chamberland (2015), pp. 106–108), $f(x, y) = -(x^2 - 1)^2 - (x^2 y - x - 1)^2$ has only two critical points, one at $(-1, 0)$ and one at $(1, 2)$, both global maxima.

$(\lambda(\hat{\theta}_h), \hat{\theta}_h)$ agree. Given that $\lambda(\theta_h)$ maximizes $r(\cdot, \theta_h)$, this follows from the Envelope Theorem. The proof again deals with kinks in v^{-n} at either θ_h or $\lambda(\theta_h)$.

Using Lemmas 17 and 19, Lemma 20 and Corollary 4 show that ψ has a unique maximum on the interval D, that is, as one hikes northward along the path λ. This uses the concavity properties established for r, with the usual complexities at kink points. Finally, Lemma 21 shows that if θ_h^* is the unique maximizer of ψ, 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 $(\tilde{\theta}_T, 1)$. We can then argue as before that any optimum of r occurs either in Θ, 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, and then proceed as above, but exchange the roles of θ_l and θ_h, so that one defines $\tilde{\lambda}(\theta_l)$ by first maximizing along north-south slices where θ_l is held constant, and then hikes eastward along the path defined by $\tilde{\lambda}$.

6.2 Sufficiency and Existence

The sufficiency part of Theorem 2 follows intuitively since any profile satisfying *PS*, *IO*, and *OB* corresponds to a critical point of r, and so if $C1$ and $C2$ held, would be a best-response for each n by Proposition 1. We show how to modify strategies in an inessential way outside of the interval served by each firm so that this holds. 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 inessential ways 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 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 (as mentioned, this is less economically natural). For simplicity, we assume the agent has a constant outside option $\bar{u} > -\infty$, and that the set of actions is $[0, \bar{a}]$.[36] It is direct that Proposition 5 (Online Appendix) continues to hold with mixed strategies, so that cross-subsidization is not beneficial. So, we will restrict each firm to action-surplus pairs that do not strictly lose money if accepted. This rules out the nonsensical equilibrium discussed at the end of Section 4.1.1. We will show existence in the game where firms can randomize over such strategies.

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

To formalize, for any convex function j, let $G(\theta, j) \subseteq [0, \bar{a}]$ be the subdifferential of j at θ. 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 θ without attracting another type.[37] Let $\rho^n \equiv \min\left\{\bar{u} - 1, \min_{a\in[0,\bar{a}]} V^n(a)\right\} > \infty$, noting that there is no benefit to being able to offer surplus below ρ^n. Let

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

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

Since v^n is convex, $G(\cdot, v^n)$ is a singleton almost everywhere, and 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 π^n, and where φ_e^n is the efficient tie-breaking rule.[38] 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) $(\bar{W}^n, \bar{\Pi}_e^n)_{n=1}^N$ *has an equilibrium.*

The proof (see Appendix D) uses Reny (1999), Corollary 5.2. The novel part of the proof, which may be of more general applicability, deals with the fact that a strategy "near" μ^{-n} can with small probability be far from the support of μ^{-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 θ 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.

[37]The max operator is valid, since, as the Appendix establishes, $G(\cdot, \cdot)$ is upper hemicontinuous.

[38]This reduction could have been made earlier, but it was convenient to make the action schedules explicitly.

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

9 Appendix A: Proofs for Section 4.1

We show that each of the asserted properties in Theorem 1 hold (see also the online appendix).

9.1 Proofs 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.[39] Thus, for ε sufficiently small but positive, offering all types surplus $v_*^n(\theta) - \varepsilon$ and action $\alpha_*^n(\theta)$ earns at least ε on a positive-measure set of types, and hence, n must earn strictly positive profits in equilibrium. □

9.2 Proofs for Section 4.1.2

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

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,

[39] If n offers $v^{-n}(\theta)$, then firms other than n win with positive probability since ties are broken equiprobably.

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 $\pi^n_v = -1$, 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 on a full-measure set of θ 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,

$$\Pi(s^n, s^{-n}) = \int \pi^n(\theta, s^n)\varphi^n(\theta, s)h(\theta)d\theta \leq \int_{P^{O,n}} \pi^n(\theta, s^n)\varphi^n(\theta, s)h(\theta)d\theta$$
$$= \int_{P^{O,n}} \pi^n(\theta, s^O)\varphi^n(\theta, s)h(\theta)d\theta.$$

Combining these two inequalities,

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

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

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

But, at any given θ, $\varphi^{n'}(\theta, s) + \varphi^{n''}(\theta, s) \leq 1$, and so the 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. □

9.3 Proof for Section 4.1.3

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

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

Assume that $\theta^{n'}_{\text{inf}} < \theta^n_{\text{sup}}$. Conditional on $\varphi^{n'}(\theta, s) > 0$, with probability one $\pi^{n'}(\theta, \alpha^{n'}, v^{n'}) \geq 0$ by *PP*, and $\pi^n(\theta, \alpha^{n'}, v^{n'}) \leq 0$ by *NP*. Hence, for any $\varepsilon \in (0, (\theta^n_{\text{sup}} - \theta^{n'}_{\text{inf}})/2)$ 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,

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

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 ε 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 θ_h^n or below θ_l^n either loses for sure or ties but earns 0. □

9.4 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 π 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 θ_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 ε 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 one 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 π is strictly concave in a, this perturbation is strictly profitable for sufficiently small ε and q, contradicting the optimality of $\hat{s}^n$. See the online appendix for details.

Corollary 1 *Every Nash Equilibrium that satisfies NEO has PS.*

Proof Assume that for some $n' > n$, and for some $\hat{\theta} \in (\theta_l, \theta_h)$, $v^{n'} = v^n$. 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 $[\hat{\theta}, \theta_h]$, contradicting that n wins with probability one conditional on $\theta \in (\theta_l, \theta_h)$. □

9.5 Proofs for Section 4.1.5

Lemma 5 *Let* $\kappa \in [0,1]$*. Then,* $((\kappa - H)/h)_\theta = -1 - (\kappa - H)\, h'/h^2 < 0$.

Proof The equality is immediate. 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. □

10 Appendix B: Proofs for Section 5

10.1 Proofs for Section 5.1

Proof of Theorem 3 Let us start from Claim 2. 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 $\{v_*^{n'}\}_{n' \in \{1,\ldots,N\}}$. But, θ 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 non-empty, then v^n must cross $v_*^{(2)}$ exactly twice, once where $v_*^{(2)}$ is shallow, and once where it is steep. J^n is thus an interval, where if $n \notin \{1, N\}$, then J^n is of the form $\left(\underline{J}^n, \bar{J}^n\right)$, if J^1 is non-empty, it is of the form $[0, \bar{J}^1)$, and if J^N is non-empty, it is of the form $(\underline{J}^N, 1]$.

Let us turn to Claim 1. On $\left(\bar{J}^n, \underline{J}^{n+1}\right)$, the agent by definition strictly prefers complete information. Since by Claim 2, 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 $\theta_*^n \in \left[\bar{J}^n, \underline{J}^{n+1}\right]$. But, $v_*^{(2)}(\theta_*^n) = v_*^{(1)}(\theta_*^n) > v^O(\theta_*^n)$, where $v^O(\theta)$ is the surplus of type θ with incomplete information, since each firm makes strictly positive profits on all types in equilibrium. Hence, $\theta_*^n \in \left(\bar{J}^n, \underline{J}^{n+1}\right)$, which is thus non-empty. Finally, let θ^n be the boundary between firms n and $n+1$ under incomplete information. Then, $\theta^n \in \left[\bar{J}^n, \underline{J}^{n+1}\right]$, 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 $\theta^n \in \left(\bar{J}^n, \underline{J}^{n+1}\right)$. Finally, note that at θ_*^n, both firms earn strictly positive profits under incomplete information, but zero under complete information. $\square$

10.2 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

$$\begin{aligned} q\left(k^n, k^{n+1}, \theta\right) &= V^n\left(\gamma^n(\theta, k^n)\right) + \theta\gamma^n(\theta, k^n) - \left(V^{n+1}\left(\gamma^{n+1}(\theta, k^{n+1})\right) + \theta\gamma^{n+1}(\theta, k^{n+1})\right) \\ &\quad + \frac{k^n - H(\theta)}{h(\theta)}\left(\gamma^{n+1}(\theta, k^{n+1}) - \gamma^n(\theta, k^n)\right). \end{aligned}$$

Note that by the definition of γ^n, $V_a^n\left(\gamma^n(\theta, k^n)\right) + \theta = \left(k^n - H(\theta)\right)/h(\theta)$, and so $q_\theta(k, k, \theta) = ((k - H(\theta)/h(\theta))_\theta - 1)\left(\gamma^{n+1}(\theta, k) - \gamma^n(\theta, k)\right) < 0$. 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 κ but not on $\bar{u}$. Thus, given κ^M, and given that the merged firm chooses to serve type θ, it does so optimally with Firm n_l for θ below θ^{M,n_l}, Firm $n \in \{n_l, \ldots, n_h - 1\}$ for θ between $\theta^{M,n-1}$ and $\theta^{M,n}$, and Firm n_h for θ above θ^{M,n_h-1}.

Lemma 6 *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

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

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

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

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

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

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

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 $[\underline{\theta},1]\subseteq[0,1]$ if there are $\underline{\theta}\leq\tau_l\leq\tau_h\leq 1$ such that g is strictly increasing on $[0,\tau_l)$, constant on $[\tau_l,\tau_h]$, and strictly decreasing on $(\tau_h,1]$, where at least one of $\tau_l>0$ or $\tau_h<1$ holds.

Lemma 7 *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 decreasing, and if $\kappa^M\leq\kappa^{n_l}$, then v^O-v^M is increasing.*

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 κ 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 $n^*\in\{n_l,\ldots,n_h-1\}$ (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 6, we have $\kappa^M>\kappa^{n^*}$. Further, for any $n\leq n^*-1$ (if there are any such), we have $\theta^{M,n}>\theta^{O,n^*}$, again by Lemma 6. Thus, for any $\theta\in[\theta_l,\theta^{M,n^*})$, we have that the active firm in the monopoly at θ 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, $v^O_\theta>v^M_\theta$ 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 non-empty.

Note next that for $n\geq n^*+1$, since $\kappa^M\leq\kappa^{n^*+1}\leq\kappa^n$, by Lemma 6, $\theta^{M,n}>\theta^{O,n}$. On $(\theta^{M,n^*},\theta^{M,n^*+1})$ both the oligopoly and the monopolist are using Firm n^*+1, but $v^O_\theta\leq v^M_\theta$ 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 the oligopoly, and is acting according to a strictly lower κ, since the relevant κ in oligopoly is at least $\kappa^{n^*+2}>\kappa^{n^*+1}\geq\kappa^M$. Hence, $v^O_\theta<v^M_\theta$ on $(\theta^{M,n^*+1},\theta_h]$.

Assume next that $\theta^{M,n^*}\leq\theta^{O,n^*}$. As above, for $n\leq n^*-1$, by Lemma 6, $\theta^{M,n}>\theta^{O,n}$, 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, $v^O_\theta>v^M_\theta$. On the interval $\theta\in\left(\theta^{M,n^*-1},\theta^{M,n^*}\right)$ both the oligopoly and the monopoly are using firm n^*, and so, since $\kappa^M\geq\kappa^{n^*}$, $v^O_\theta\geq v^M_\theta$. Finally, by Lemma 6 $\kappa^M<\kappa^{n^*+1}$, 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 non-empty.

Consider next $\kappa^M \leq \kappa^{n_l}$. By Lemma 6, $\theta^{M,n} < \theta^{O,n}$ 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 non-empty interval $\left[\theta^{M,n_l}, \theta_h\right]$. Similarly if $\kappa^M \geq \kappa^{n_h}$. then, by Lemma 6, $\theta^{M,n} > \theta^{O,n}$ for all $n \in \{n_l, \ldots, n_h - 1\}$, and so $v_\theta^O > v_\theta^M$ everywhere on $[\theta_l, \theta_h]$, and strictly so on the non-empty interval $\left[\theta_l, \theta^{M,n_h - 1}\right]$. $\square$

Proof of Theorem 5 By Lemma 7, $v^O - \hat{v}^M$ 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 non-empty, and hence $\hat{\kappa}^M \in (\kappa^{n_l}, \kappa^{n_h})$. $\square$

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 θ_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 $n \in \{n_l', \ldots, n_h' - 1\}$, $\theta^{M,n} > \theta^{O,n}$. 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 θ^* be the lower boundary of firm n_l' in oligopoly. If $n_l' = n_l$, let $\tilde{u}(\theta) = \bar{u}(\theta)$ for all θ. If $n_l' > n_l$, then let $\tilde{u} = \bar{u}$ for $\theta < \theta_l$, and define $\tilde{u} = \max_{n \in \{n_l, \ldots, n_l' - 1\}} v^n$ for $\theta > \theta_l$. As a maximum of convex functions, $\tilde{u}$ is convex, and by stacking, $\tilde{u}' < \gamma^{n_l'}$ for any θ and $\kappa \in [0,1]$. Thus, from Firm n_l''s optimal boundary condition at the bottom

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

Note that $\omega_\theta^{n_l'}\left(\theta, \kappa^{n_l'}\right) = ((\kappa^{n_l'} - H/h)_\theta - 1)(\tilde{u}' - \gamma^{n_l'}) + ((\kappa^{n_l'} - H)/h)\tilde{u}'' > 0$ for all $\theta \leq \theta^*$, using that $((\kappa^{n_l'} - H)/h)_\theta - 1 < 0$, $\tilde{u}' - \gamma^{n_l'} < 0$, $\kappa^{n_l'} - H(\theta) \geq \kappa^{n_l'} - H(\theta^*) > 0$, 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 θ_l^M is optimal.

Hence, we must have $\kappa^M < \kappa^{n_h'}$. Now let θ^* 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 $\tilde{u} = \max_{n \in \{n_h' + 1, \ldots, n_h\}} v^n$ for $\theta < \theta_h$. Firm n_h''s optimal boundary condition at θ^* in the oligopoly is

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

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

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

since $((\kappa^{n'_h} - H)/h)_\theta - 1 < 0$, $\bar{u}' - \gamma^{n'_h} > 0$, $\kappa^{n'_h} - H(\theta) \leq \kappa^{n'_h} - H(\theta^*) < 0$, by OB, and $\tilde{u}'' \geq 0$. Hence, since $0 = \omega^{n'_h}\left(\theta^*, \kappa^{n'_h}\right)$, we have $0 \geq \omega^{n'_h}\left(\theta_h^M, \kappa^{n'_h}\right)$. 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 $0 > \omega^{n'_h}\left(\theta_h^M, \kappa^M\right)$. This contradicts that the merged firm has set θ_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 $\left[\theta_l^M, \theta_h^M\right]$ be the span of the merged firm, with associated v^M and κ^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,

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

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

11 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 $C1$ and $C2$. 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 θ_h does not make sense. But, from (5), since π is strictly concave in a, we have that $0 = \pi_a(\theta_h, \alpha(\theta_h), v)\,(a^{-n}(\theta_h) - \alpha(\theta_h)) + \pi(\theta_h, \alpha, v) > \pi(\theta_h, a^{-n}(\theta_h), v) - \pi(\theta_h, \alpha, v) + \pi(\theta_h, \alpha, v) = \pi(\theta_h, a^{-n}, v) = \pi(\theta_h, a^{-n}, v^{-n})$. □

11.1 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 θ_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 $\tilde{\kappa}(\theta_l, \theta_h) = H(\theta_l)$. □

Lemma 8 *The function r has a maximum, and at any maximum (θ_l, θ_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 (θ_l, θ_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 (θ_l, θ_h), γ is continuous in κ, $\tilde{v}$ 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, using Lemma 5, $\tilde{\alpha}$ 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 (θ_l, θ_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 (θ_l, θ_h) is optimal, implies $\theta_h = 1$. To do so, note first that for all θ_h' on a neighborhood of θ_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 20). 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 θ_h' on a neighborhood of θ_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 (θ_l, θ_h) is a maximum of r, and so $r(\theta_l, \theta_h) > 0$ by $C2$, 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 (θ_l, θ_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 (θ_l, θ_h), using stacking and $C1$. Part (iii) follows immediately, with the equality of payoffs following as the relevant domains of integration agree. □

Lemma 9 *There exists $(\underline{m}, \overline{m}) \ni \theta^x$ 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 $C1$ 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 θ^x. Formally, by stacking and $C1$, for $\theta > \theta^x$, $a^{-n}(\theta) > \gamma(\theta, 0) \geq \gamma(\theta, H(\theta)) = \alpha_*(\theta)$, 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 $[\theta^x, 1]$, and so single-crosses 0 from above at most once on $[\theta^x, 1]$. 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 $\pi(\theta^x, a^{-n}, v^{-n}) < 0$, take $\bar{m} = \theta^x$. Construct $\underline{m}$ similarly. □

Strategy s^n is *dominant* on (τ_l, τ_h) if (τ_l, τ_h) is a maximal interval such that $v^n > v^{-n}$.

Lemma 10 *Let (α, v) be any feasible menu for n, let v be dominant on (τ_l, τ_h), and let $\pi(\cdot, \alpha, v) \geq 0$ on (τ_l, τ_h). Then, $(\tau_l, \tau_h) \cap [\underline{m}, \overline{m}] \neq \varnothing$.*

Proof Let $\tau_l \geq \bar{m} \geq \theta^x$, 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 $\tau > \overline{m} \geq \theta^x$, and using $C1$, it follows that $a^{-n}(\tau) > \alpha_*(\tau)$, 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. □

Lemma 11 *Assume stacking. Fix n and s^{-n} satisfying $C1$ and $C2$. Then, for each $\hat{s}$ there is (θ_l, θ_h) with $\Pi(\hat{s}, s^{-n}) \leq r(\theta_l, \theta_h)$.*

Proof Intuitively, let $\overline{m}^* \geq \overline{m}$ capture any region of dominance of v that contains $\overline{m}$, 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 $[\underline{m}^*, \overline{m}^*]$, and adding any agent in $(\underline{m}, \overline{m})$ that it does not already serve with probability one. But, $\tilde{s}(\underline{m}^*, \overline{m}^*)$ accomplishes exactly this, and does so optimally in the relaxed problem, and hence its associated payoff $r(\underline{m}^*, \overline{m}^*)$ is at least as high as $\Pi(s^n, s^{-n})$.

To formalize this, note that using Proposition 5, we can wlog assume that (α, 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 $\theta^x \leq \underline{I}_H \leq \overline{m} \leq \bar{I}_H$. In this case, define $\overline{m}^* = \bar{I}_H$. If there is no such interval, define $\overline{m}^* = \overline{m}$. Similarly, if v dominates v^{-n} on an interval I_L with $\underline{I}_L \leq \underline{m} \leq \bar{I}_L \leq \theta^x$, then define $\underline{m}^* = \underline{I}_L$, and if there is no such interval, define $\underline{m}^* = \underline{m}$.

Consider first any positive-lengthed interval $J \subseteq [\overline{m}^*, 1]$ 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 $\overline{m}^* \geq \overline{m}$, $\pi(\theta, \alpha, v) < 0$ for all $\theta > \overline{m}^*$. Hence, excluding J from the domain of the integral in Π increases its value.

By Lemma 10, and since we have wlog taken (α, v) to strictly lose money nowhere, there is no positive -lengthed interval $J = (\underline{J}, \bar{J})$ with $\underline{J} \geq \overline{m}^*$ or and $\bar{J} < \underline{m}^*$ on which v is dominant. We thus have $\Pi(s) \leq \int_{\underline{m}^*}^{\overline{m}^*} \pi(\theta, \alpha, v)\varphi(\theta, s)h(\theta)d\theta$. Define $\hat{v} = \max(v, v^{-n})$, with associated $\hat{\alpha}$, where at all θ where $v(\theta) \geq v^{-n}(\theta)$, we can take $\hat{\alpha} = \alpha$, and at almost all θ 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 $\theta \in (\underline{m}^*, \overline{m}^*)$ such that $\varphi(\theta, s) < 1$. Since by construction, φ is 1 on I_H and I_L (if these sets exist), it follows that $\theta \in [\underline{m}, \overline{m}]$. 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 $\Pi(s) \leq \int_{\underline{m}^*}^{\overline{m}^*} \pi(\theta, \hat{\alpha}, \hat{v}) h(\theta) d\theta$. But, $\hat{v} \geq v^{-n}$ by construction, and so (2) and (3) are satisfied in $\mathcal{P}(\underline{m}^*, \overline{m}^*)$, 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 $\Pi(s) \leq r(\underline{m}^*, \overline{m}^*)$. □

Proof of Proposition 1 Immediate from Lemmas 8 and 11, as discussed in main text. □

11.2 Proofs for Section 6.1

Lemma 12 *Any maximum of r is in $R = [0, \theta^x] \times [\theta^x, 1]$.*

Proof Let (θ_l, θ_h) with $\theta_h < \theta^x$, be a maximum of r. Then, $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$ by Lemma 8, and so, since $\tilde{\kappa} \leq H(\theta_h) < H(\theta^x) \leq 1$, it follows from stacking and the definition of θ^x that $a^{-n} < \gamma(\cdot, \tilde{\kappa})$ for $\theta < \theta^x$, and so v crosses v^{-n} from *below* at θ_h. But, by Lemma 8, (*iii*), v is single-dominant on (θ_l, θ_h), a contradiction. Thus, $\theta_h \geq \theta^x$. Similarly, $\theta_l \leq \theta^x$. □

11.2.1 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 13 *Considered as a function on $\tilde{R}$, r is continuously differentiable, with*

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

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

Proof The right side of (12) has the same form as (5). As in the analysis of *OB* in Section 1.4, this is the value of increasing θ_h by increasing the action immediately to the left of θ_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).[40] On $\tilde{R}$, r_{θ_h} and r_{θ_l} are continuous. Hence, r is continuously differentiable. □

[40] 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 23 in the online appendix), and then differentiate with respect to θ_h and manipulate.

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

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

where the second equality uses L'Hospital's rule and third uses continuity of r_{θ_h} on (ι'_l,ι'_h).

Recall that Θ 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 14 *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 (θ_l,θ_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 θ_l. See the online appendix for details.

11.2.2 Essentially Unique Optimality

Fix a function $f : [x_l,x_h] \to \mathbb{R}$, 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 $f^-_x(x)f^+_x(x) \le 0$, 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 $f_x\,(x_l) \equiv f^+_x\,(x_l) \le 0$, and that x_h is a *critical point* of f if $f_x\,(x_h) \equiv f^-_x\,(x_h) \ge 0$. Any maximum of f is at a critical point.

Note that ι is continuously differentiable on each $\tilde{R}$, and is continuous on R. Hence, since ι_{θ_l}, ι_{θ_h}, and ι_κ are strictly positive, and since $\iota\,(\theta^x,\theta^x,\cdot) = 0$, L_N is continuous, strictly decreasing, and goes through (θ^x,θ^x). The locus $L_S < L_N$ has the same properties.

Lemma 15 *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, $r_{\theta_l}(\theta_l,\theta_h) = -\pi(\theta_l,\gamma(\cdot,\tilde{s}\,(\theta_l,\theta_h))h(\theta_l)$, and if $r(\theta_l,\theta_h) > 0$, then $r_{\theta_l}(\theta_l,\theta_h) < 0$.*

Proof Fix (θ_l,θ_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 13, and again exploiting that above L_S, $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$,

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

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 γ 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 γ. 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 (θ_l, θ_h) above L_N is similar. □

Assumption 1 (L_S hits the western boundary of R) $\iota\,(0, 1, 1) \geq 0$.

Note in particular that since ι 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 θ_T by $\iota(0, \theta_T, 0) = 0$ if there is such a $\theta_T \geq \theta^x$, and by $\theta_T = 1$ otherwise. This is the latitude at which L_N exits R. For Firm 1, $\theta^x = 0$, and hence $\theta_T = 0$. Let $A = \{(0, \theta_h) | \theta_h \geq \theta_T\}$.

Corollary 3 *Any maximum of r occurs in $(\Theta \cup A) \backslash (\theta^x, \theta^x)$.*

Proof Since r is strictly positive at an optimum, then below L_S, $r_{\theta_h} > 0$ by Lemma 15, 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, $r\,(\theta^x, \theta^x) = 0$, and so (θ^x, θ^x) is not an optimum either. □

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 16 *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 ψ is continuous on $[\theta^x, 1]$. On D, λ 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 θ_l is a critical point, then $r_{\theta_l} = 0$ and so by Lemma 14, $r_{\theta_l \theta_l} < 0$. Thus θ_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 θ_l is a critical point. Then, since $\pi_a(\theta_l, \gamma(\theta_l, \tilde{\kappa}), \tilde{v}) \geq 0$ and since a^{-n} jumps upwards at θ_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 θ_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 θ_l' in a neighborhood to the left of θ_l. If instead $r^-_{\theta_l} = 0$, then, by 14 applied to the rectangle $\tilde{R}$ to the left of (θ_l, θ_h), $r(\cdot, \theta_h)$ is strictly concave on a neighborhood to the left of θ_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 θ'_l in a neighborhood to the left of θ_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 θ_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 θ_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, ψ is continuous, and the set of maximizers of $r(\cdot, \theta_h)$ is upper hemicontinuous in θ_h. But then, since λ is single-valued on D by the first part of the proof, it is continuous as a function on D. □

Lemma 17 *The set $D = \{\theta_h > \theta^x | \psi(\theta_h) > 0\}$ is an interval of the form $(\theta^x, \bar{D})$.*

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 $(\theta^x, \hat{\theta}_h) \subseteq D$. 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 $\theta'_h \in (\theta^x, \hat{\theta}_h)$, a parallel shift downwards of $\hat{v}$ wins a positive interval of types (θ'_l, θ'_h) strictly profitably, and hence $\psi(\theta'_h) > 0$. Note first that PP (and in particular Proposition 5 in the online appendix) 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 $C1$ and stacking, $\hat{v}$ is strictly shallower than v^{-n} above θ^x. So, $\delta > 0$, and since $\hat{v} - \delta$ is strictly steeper than v^{-n} below θ^x, it follows that s' is single-dominant on (θ'_l, θ'_h) for some $\theta'_l < \theta^x$, 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 $(\theta'_l, \theta'_h) \supseteq (\theta^x, \theta'_h)$. Hence, $\psi(\theta'_h) \geq r(\theta'_l, \theta'_h) > 0$, and so $\theta'_h \in D$. □

Lemma 18 *Let $(\lambda(\theta_h), \theta_h) \in L_S$ with $\theta_h \in D$. Then, ψ is strictly increasing at θ_h, and so θ_h is not a critical point of ψ.*

Proof The basic idea is that by Lemma 15, $r_{\theta_h}(\theta_l, \theta_h) > 0$ anywhere near L_S. But, since L_S is decreasing, as one moves a little above L_S, the constraint on θ_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 $\theta_h > \theta^x$. 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 13, r_{θ_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, and so, by examination of (12) and by Lemma 15, it follows that r_{θ_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 15, $r_{\theta_h}(\theta_l, \theta_h) > 0$.

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

$$\psi^+_{\theta_h}(\theta'_h) = \lim_{\varepsilon \downarrow 0} \frac{\psi(\theta'_h + \varepsilon) - \psi(\theta'_h)}{\varepsilon} \geq \lim_{\varepsilon \downarrow 0} \frac{r(\lambda(\theta'_h), \theta'_h + \varepsilon) - r(\lambda(\theta'_h), \theta'_h)}{\varepsilon} = r_{\theta_h}(\lambda(\theta'_h), \theta'_h), \quad (15)$$

where the inequality follows since for small ε, $\lambda(\theta'_h)$ is feasible at $\theta'_h + \varepsilon$.[41] Thus, since $r_{\theta_h}(\lambda(\theta_h), \theta_h) > 0$ by Lemma 15, $\psi^+_{\theta_h}(\theta_h) > 0$.

Finally, let us show that $\psi^-_{\theta_h}(\theta_h) > 0$. Let $\rho = r_{\theta_h}(\lambda(\theta_h), \theta_h)/2$. Since r_{θ_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 τ by the definition of a right derivative. Thus, $j(\theta_h) \geq 0$, and we are done. □

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

Lemma 19 *On $D' \backslash K$, ψ is continuously differentiable, with $\psi_{\theta_h}(\theta_h) = r_{\theta_h}(\lambda(\theta_h), \theta_h)$. For all $\theta_h \in D' \cap K$, ψ 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 non-differentiable 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 15, $r_{\theta_l}(\theta_l, \theta_h) < 0$ 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 differentiable on any given $\tilde{R}$, ψ is also continuously differentiable, with

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

since λ 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 $\bar{J}(\theta_l) = \max\{\theta_h | \lambda(\theta_h) = \theta_l\}$.[42] Since λ is constant on $J(\theta_l)$, if $J(\theta_l)$ is non-empty, then for all $\theta_h \in J(\theta_l) \backslash K_2$, we have again have that ψ is continuously differentiable and (16).

Case 3 Consider next $\theta_h \in (\{\underline{J}(\theta_l)\}_{\theta_l \in K_1} \cup \{\bar{J}(\theta_l)\}_{\theta_l \in K_1}) \backslash K_2$. 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 θ_h, $\theta'_h \notin K_2$,

[41] 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.

[42] These correspond to the bottoms and tops of the vertical segments of the path in Figure 5.

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 θ_h. Finally, if $(\underline{J}(\theta_l), \bar{J}(\theta_l))$ is non-empty, then $\psi_{\theta_h}(\theta_h') = r_{\theta_h}(\lambda(\theta_h'), \theta_h')$ for θ_h' on a neighborhood above θ_h by Case 2. Now, note by (12), that r_{θ_h} does not depend on $a^{-n}(\theta_l)$, and so r_{θ_h} is continuous on $D'\backslash K_2$, even though $\lambda(\theta_h) \in K_1$.[43] But then, by continuity of λ, $\psi_{\theta_h}(\cdot)$ is continuously differentiable at θ_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 θ_h, ψ is continuously differentiable with $\psi_{\theta_h} = r_{\theta_h}$ by the previous cases, and λ is continuous, and so

$$\psi^+_{\theta_h}(\theta_h) = \lim_{\varepsilon \downarrow 0} \frac{\psi(\theta_h + \varepsilon) - \psi(\theta_h)}{\varepsilon} = \lim_{\varepsilon \downarrow 0} \psi_{\theta_h}(\theta_h + \varepsilon) = \lim_{\varepsilon \downarrow 0} r_{\theta_h}(\lambda(\theta_h + \varepsilon), \theta_h + \varepsilon) = r^+_{\theta_h}(\lambda(\theta_h), \theta_h),$$

using L'Hospital's rule at the second inequality, and where the last equality uses from the last part of Case 3, r_{θ_h} is continuous on $D'\backslash K_2$. Similarly, $\psi^-_{\theta_h}(\theta_h) = r^-_{\theta_h}(\lambda(\theta_h), \theta_h)$. □

Lemma 20 *Every critical point θ_h of ψ on D is a strict local maximum of ψ. That is, for all $\theta_h' \neq \theta_h$ in some neighborhood of θ_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 18, any critical point of ψ on D is an element of D', except in the case that $\theta_h = 1 \in D\backslash D'$, in which case we are trivially done by Lemma 18. But, on D', Lemma 19 lets us relate the local concavity properties of ψ to those we establish for r in Lemma 14. We go through the same cases as in Lemma 19.

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 14 applies and $r_{\theta_l\theta_l} < 0$, and thus, by the Implicit Function Theorem, $\lambda_{\theta_h} = -r_{\theta_l\theta_h}/r_{\theta_l\theta_l}$. Since (16) holds on a neighborhood of θ_h,

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

Since θ_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 14 $r_{\theta_l\theta_l} r_{\theta_h\theta_h} - r^2_{\theta_l\theta_h} > 0$. Hence, $\psi_{\theta_h\theta_h}(\theta_h) < 0$, and θ_h is a strict local maximum of ψ.

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 $J(\theta_l)\backslash K_2$ is open, by Case 2 of Lemma 19, (16) holds on a neighborhood of θ_h, and so, since λ 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

[43] Note that r_{θ_h} depends on θ_l only through $\tilde{\kappa}$, and that $\tilde{\kappa}$ does not depend on $a^{-n}(\theta_l)$.

$(\lambda(\theta_h), \theta_h) \in \Theta$, then by Lemma 14, if $\psi_{\theta_h}(\theta_h) = 0$, then $\psi_{\theta_h\theta_h}(\theta_h) < 0$, so θ_h is a strict local maximum of ψ. 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 14 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, θ_h is a strict local maximum of ψ.

Case 3 Consider next $\theta_h \in (\{\underline{J}(\theta_l)\}_{\theta_l \in K_1} \cup \{\bar{J}(\theta_l)\}_{\theta_l \in K_1}) \backslash K$. 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, ψ is strictly concave on a neighborhood just above θ_h, while by Case 1, ψ is strictly concave on a neighborhood just below θ_h. Hence, again, θ_h is a strict local maximum of ψ.

Case 4 Finally, consider $\theta_h \in K_2 = K \cap D'$. Since $\tilde{\kappa} \in [0, 1]$, and since $\theta_h > \theta^x$, we have that $a^{-n} - \gamma$ is positive and bounded away from 0 and ∞ on a neighborhood of θ_h by stacking and $C1$. At any point θ'_h of continuity of a^{-n}, and repeating (12) for convenience,

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

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 π and π_a. Thus any discontinuity in ψ_{θ_h} at θ_h is driven by an upward jump of a^{-n} at θ_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 θ_h is not a critical point. If $\pi > 0$, then since a^{-n} jumps up at θ_h, we have $\psi^-_{\theta_h} > \psi^+_{\theta_h}$. Assume that θ_h is a critical point, so that $\psi^-_{\theta_h}\psi^+_{\theta_h} \leq 0$. If $\psi^-_{\theta_h} > 0 > \psi^+_{\theta_h}$, then θ_h is a strict local maximum of ψ. 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 θ on a neighborhood to the right of θ_h. Similarly if $\psi^-_{\theta_h} = 0$, then $\psi^+_{\theta_h} < 0$, and $\psi_{\theta_h\theta_h} < 0$ for all θ on a neighborhood to the left of θ_h. In each case θ_h is again a strict local maximum of ψ. □

Corollary 4 *There is a unique critical point of ψ on $[\theta^x, 1]$, and it uniquely maximizes ψ.*

Proof By Weierstrass' Theorem, ψ has a maximum on $[\theta^x, 1]$. But, for any maximizer, θ^*_h, $\theta^*_h \neq \theta^x$, since $\Theta(\theta^x) = \{\theta^x\}$, and so $\psi(\theta^x) = r(\theta^x, \theta^x) = 0$. If $\theta^*_h \in (\theta^x, 1)$, then $\psi^-_{\theta_h}(\theta^*_h) \geq 0$, and $\psi^+_{\theta_h}(\theta^*_h) \leq 0$, since θ^*_h is a maximizer. Similarly, if $\theta^*_h = 1$, then $\psi^-_{\theta_h}(\theta^*_h) \geq 0$. In each case, θ^*_h is critical by definition. Thus, since any maximum is a critical point, ψ has a critical point.

Let θ^*_h be any critical point of ψ, and let us show that θ^*_h is the unique maximizer of ψ (and hence θ^*_h is the unique critical point of ψ). Wlog, assume that for some $\theta^{**}_h > \theta^*_h$, $\psi(\theta^{**}_h) \geq \psi(\theta^*_h)$. Then ψ attains a minimum $\theta^{\min}_h$ on the compact set $[\theta^*_h, \theta^{**}_h]$. But, since θ^*_h is a strict local maximum of ψ, we have $\psi\left(\theta^{\min}_h\right) < \psi(\theta^*_h) \leq \psi(\theta^{**}_h)$, and so $\theta^{\min}_h \in (\theta^*_h, \theta^{**}_h)$. Hence, since $\theta^{\min}_h$ is an interior minimum, $\psi^-_{\theta_h}\left(\theta^{\min}_h\right) \leq 0$, and $\psi^+_{\theta_h}(\theta^{\min}_h) \geq 0$, and so $\theta^{\min}_h$ is a critical point of ψ. But then, by Lemma 20, $\theta^{\min}_h$ is a strict local maximum, a contradiction. □

Lemma 21 *Let θ_h^* be the unique maximizer of ψ. 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 non-empty, $r(\theta_l^{**}, \theta_h^{**}) > 0$, 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 16, $\theta_l^{**} = \lambda(\theta_h^{**})$. 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 19, $\psi_{\theta_h}^{+}(\theta_h^{**}) \leq 0$ and $\psi_{\theta_h}^{-}(\theta_h^{**}) \geq 0$, and so by Corollary 4 $\theta_h^{**} = \theta_h^*$, and we are done. $\square$

11.3 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 $C1$ on $[\theta_l, \theta_h]$. But then, by *IO*, if $n < N$, then $\pi_a(\theta_h, \hat{\alpha}, \hat{v}) < 0$, and by $C1$ 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 θ_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 $C2$ on $[\theta_l, \theta_h]$.

Let us re-define $(\hat{\alpha}, \hat{v})$ outside of $[\theta_l, \theta_h]$ to satisfy $C1$ and $C2$ 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 θ. That is, actions and surplus modified outside of $[\theta_l, \theta_h]$ to ensure that $C1$ 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 θ_h, and similarly at θ_l. By stacking, (α, v) is single-dominant on (θ_l, θ_h), and so, since (α, v) and $(\hat{\alpha}, \hat{v})$ agree on $[\theta_l, \theta_h]$, (α, v) and $(\hat{\alpha}, \hat{v})$ are equivalent.

To show that $C2$ holds for $\theta \notin [\theta_l, \theta_h]$, assume $(\theta_h, 1]$ is non-empty (the case $[0, \theta_l)$ non-empty 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 π 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 $C2$ 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 α 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 16. 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 19. Finally, since by Corollary 4, ψ is strictly single-peaked on the interval D, $\theta_h = \theta_h^*$ by Lemma 21. 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 λ. So, we argue first that $\theta_h \in \tilde{\Theta}(\theta_l)$, second that by the analogue to Lemma 16, $\theta_h = \tilde{\lambda}(\theta_l)$, and third that by the analogue to Lemma 19 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}. □

Proof of Theorem 2: Existence We begin with three further restrictions on strategies that will not bind in equilibrium, but help us towards compactness and continuity. Recall that $BR(s^{-n}) = \arg\max_{s^n \in S^n} \Pi^n(s^n, s^{-n})$. Since $\pi_{aa}^n = V_{aa}^n$, and by definition of γ, $\gamma_\theta^n(\theta, \kappa) = \left(((\kappa - H(\theta))/h(\theta))_\theta - 1\right)/V_{aa}^n(\gamma^n(\theta, \kappa))$. 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 γ^n is continuous. But then, since h is $\mathcal{C}^1$ 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 $\gamma^n(\theta, \kappa) \le \gamma^N(1,0)$ for all n, θ, and $\kappa \in [0,1]$, b_1 is a bound on the highest value and slope of any γ satisfying $C1$. We will bound the slopes of actions profiles by b_1.

C3 $0 \le \alpha^n(\theta') - \alpha^n(\theta) \le b_1(\theta' - \theta)$ for all θ, θ' with $\theta' > \theta$.

Next, let $b_2 = \min_{n, \theta_h \in [0,1], \kappa \in [0,1]} \left(\pi_a^n(\theta_h, \gamma^n(\cdot, \kappa), 0)\gamma^N(\theta_h, 0) + \pi^n(\theta_h, \gamma^n(\cdot, \kappa), 0)\right)$ 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 θ_h will be strictly positive, motivating our next restriction.

C4 $v^n(1) \ge b_2$.

For each n, let S_R^n be the subset of S^n such that $C1$–$C4$ 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 $C2$ is immediate, and that $C1$ follows because $\alpha_*^n(\theta) = \gamma^n(\theta, H(\theta))$. But then, by definition of b_1,

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

using that $\gamma_\kappa^n < 0$, and so $C3$ follows. To see $C4$, 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 π^n at surplus to the agent of 0. Thus, since b_2 is a minimum of objects of this form, $v_*^n(1) \ge b_2$. Hence, S_R^n is nonempty.

Lemma 22 *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 non-empty 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 (θ_l, θ_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 $C1$ and $C2$ are satisfied on $[\theta_l, \theta_h]$. Let (α, v) be defined from $(\hat{\alpha}, \hat{v})$ as in the proof of Theorem 2, so that as shown there, $C1$ and $C2$ are satisfied on $[0, 1]$. By stacking, and using that for $n' \neq n$, $C1$ and $C2$ are satisfied by assumption, it remains the case that (α, v) is single-dominant on $[\theta_l, \theta_h]$, and since (α, v) and $(\hat{\alpha}, \hat{v})$ agree on $[\theta_l, \theta_h]$, it follows that $(\alpha, v) \in BR(s^{-n})$. Condition $C3$ holds by construction.

To show $C4$, 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 (θ_l, θ_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 π, γ, and v yield that $\pi(\theta_h, \gamma(\cdot, \kappa), v) \geq 0$. By $C1$ and $C2$ for $n' \neq n$, and stacking, $\bar{a} - \gamma(\theta_h, \kappa) > 0$. Hence, $0 \geq \pi_a(\theta_h, \gamma(\cdot, \kappa), v)$, and so

$$0 \geq \pi_a(\theta_h, \gamma(\cdot, \kappa), v)\gamma^N(\theta_h, 0) + \pi(\theta_h, \gamma(\cdot, \kappa), v) > \pi_a(\theta_h, \gamma(\cdot, \kappa), 0)\gamma^N(\theta_h, 0) + \pi(\theta_h, \gamma(\cdot, \kappa), 0) - b_2 \geq 0,$$

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 (α, v) is a best response and satisfies $C1 - C4$, we are done. $\square$

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 $(S^n_R, \Pi^n)_{n=1}^N$ has a pure-strategy equilibrium, since by Lemma 22, $BR^n(s^{-n}) \cap S^n_R$ is nonempty, and so in a Nash equilibrium of $(S^n_R, \Pi^n)_{n=1}^N$, 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^n_R, 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 (non-empty) set S^n_R is convex and compact. To prove convexity, let (α^n_1, v^n_1) and $(\alpha^n_2, v^n_2) \in S^n_R$, let $\delta \in [0, 1]$, and let $(\alpha^n_3, v^n_3) = (\delta\alpha^n_1 + (1 - \delta)\alpha^n_2, \delta v^n_1 + (1 - \delta)v^n_2)$. Then, (α^n_3, v^n_3) satisfies the integral condition $v^n_3(\theta) = v^n_3(0) + \int_{\underline{\theta}}^{\theta} \alpha^n_3(\tau)\, d\tau$, since integration is a linear operator, so that $(\alpha^n_3, v^n_3) \in S^n$, and it is direct that (α^n_3, v^n_3) satisfies $C1$–$C4$.

To prove compactness, let $(\alpha^{n,k}, v^{n,k})_{k=1}^\infty$ be a sequence of elements of S^n_R. By $C1$ and the definition of b_1, we have by $C3$ that for each k, $0 \leq \alpha^{n,k}(0) \leq \alpha^{n,k}(1) \leq b_1$. It follows using $C3$ and the Arzela-Ascoli Theorem (for example, Rudin (1987), Theorem 11.28, p. 245) that there exists α^n satisfying $C1$ and $C3$ and a subsequence along which $\|\alpha^{n,k} - \alpha^n\|_\infty \to 0$. Note that α^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 $C2$ and $C4$, we can, taking a further subsequence and re-indexing, assume that along the chosen subsequence $v^{n,k}(1) \to \bar{v}$, 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, $\|v^{n,k} - v^n\|_\infty \to 0$, and (ii) $(\alpha^n, v^n) \in S_R^n$. To see (i), note that for each θ and k, $v^{n,k}(\theta) = v^{n,k}(1) - \int_\theta^1 \alpha_k^n(\tau)d\tau$, and hence,

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

and so, since this is independent of θ, but converges to zero, $\|v_k^n - v^n\|_\infty \to 0$. To see (ii), note that we have checked $C1$ and $C3$, and that weak inequalities are preserved under limits, and so $C2$ and $C4$ 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 $s^k \to s$, and fix n. By stacking and since $s \in S_R$, there exist θ_l and θ_h such that $\varphi(\theta, s) = 1$ on (θ_l, θ_h) and $\varphi(\theta, s) = 0$ for $\theta \notin [\theta_l, \theta_h]$. But then, since for each n', $\|v^{n',k} - v^{n'}\| \to 0$, 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 $\|\alpha^{n,k} - \alpha^n\| \to 0$ as well, and since π is bounded and continuous, $\Pi^n(s^k) \to \Pi^n(s)$, and thus that Π^n is continuous on S_R.[44]

Fix n. Since Π^n is continuous on S_R, and since S_R^n is non-empty, 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 non-empty 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 21, $(\hat{\theta}_l, \hat{\theta}_h) = (\lambda(\theta_h^*), \theta_h^*)$. Thus, any two elements of $BR_R^n(s^{-n})$ win for sure on $(\lambda(\theta_h^*), \theta_h^*)$ and 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 non-empty, compact, convex subset of a Banach space, and that the correspondence defined by $BR_R(s) \equiv BR_R^1(s^{-1}) \times \cdots \times BR_R^N(s^{-N})$ 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. □

12 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,

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

[44] Recall that without stacking, and outside of S_R, it is easy to construct examples where payoffs are discontinuous.

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

Hence, (W, Π_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 $\mathcal{N}(\theta, v) = \left\{n | v^n(\theta) \geq v^{n'}(\theta) \, \forall n'\right\}$ is the set of firms offering maximal surplus at θ, then by efficient tie-breaking, the sum of payoffs at θ 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 θ is upper semicontinuous. Integrating across θ 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 μ^{-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 one 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 μ^n and $\varepsilon = \varepsilon'/(5+\tau)$ yields the result.

For each $v^{-n} \in W^{-n}$, let $B_\varepsilon(v^{-n})$ be the open ε-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}. Since for each $c \in C$ $W^{-n} = \cup_{\varepsilon'' \geq 0}(\partial B_{\varepsilon''}(c))$, 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 $\tilde{\partial} \equiv \cup_{c \in C} \partial B_{\hat{\varepsilon}}(c)$.

Draw the Venn diagram on $W^{-n} \backslash \tilde{\partial}$ corresponding to the set of balls $B_{\hat{\varepsilon}}(c)$ for $c \in C$, with subsets $W_1^{-n}, \ldots, W_{M^*}^{-n}$, where $M^* = 2^M - 1 < \infty$. That is, any two points in $W^{-n} \backslash \tilde{\partial}$ 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 $W^{-n} \backslash \tilde{\partial}$ and c in C are 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 $\cup_{i=1}^{M^*} W_i^{-n} = W^{-n} \backslash \tilde{\partial}$.

For $i \in \{1, \ldots, M^*\}$, 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 ε of attaining w_i. For any given θ, 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 $\hat{\mu}^{-n}$,

$$
\begin{aligned}
\bar{\Pi}_e\left(\delta_{v^n+4\varepsilon}, \hat{\mu}^{-n}\right) &= \int_{W^{-n}} \left(\int (q^n\left(\theta, v^n + 4\varepsilon\right) - v^n\left(\theta\right) - 4\varepsilon)\varphi_e^n\left(\theta, \left(v^n + 4\varepsilon, v^{-n}\right)\right) h\left(\theta\right) d\theta \right) d\hat{\mu}^{-n}\left(v^{-n}\right) \\
&\geq -4\varepsilon + \int_{W^{-n}\setminus\tilde{\partial}} \left(\int \left(q^n\left(\theta, v^n\right) - v^n\left(\theta\right)\right) \varphi_e^n\left(\theta, \left(v^n + 4\varepsilon, v^{-n}\right)\right) h\left(\theta\right) d\theta \right) d\hat{\mu}^{-n}\left(v^{-n}\right) \\
&\geq -4\varepsilon + \int_{W^{-n}\setminus\tilde{\partial}} \left(\int \left(q^n\left(\theta, v^n\right) - v^n\left(\theta\right)\right) \varphi_e^n\left(\theta, \left(v^n, v_i^{-n}\right)\right) h\left(\theta\right) d\theta \right) d\hat{\mu}^{-n}\left(v^{-n}\right) \\
&\geq -4\varepsilon + \sum_{i=1}^{M^*} \int_{W_i^{-n}} \left(w_i - \varepsilon\right) d\hat{\mu}^{-n}\left(v^{-n}\right) \\
&\geq -5\varepsilon + \sum_{i=1}^{M^*} w_i \hat{\mu}^{-n}\left(W_i^{-n}\right),
\end{aligned}
$$

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

To complete the proof, choose $\lambda > 0$ such that for all $\hat{\mu}^{-n} \in B_\lambda\left(\mu^{-n}\right)$, $\hat{\mu}^{-n}\left(W_i^{-n}\right) \geq \left(1 - \varepsilon\right)\mu^{-n}\left(W_i^{-n}\right)$ for each $1 \leq i \leq M^*$.[45] Then, for all $\hat{\mu}^{-n} \in B_\lambda\left(\mu^{-n}\right)$, and since $w_i \geq 0$,

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

where the third inequality uses $\sum_{i=1}^{M^*} \mu^{-n}\left(W_i^{-n}\right) = 1$. □

References

Aliprantis, C., and K. Border (2006): *Infinite Dimensional Analysis*. Springer Verlag, Berlin.

Armstrong, M., and J. Vickers (2001): "Competitive Price Discrimination," *Rand Journal of Economics*, pp. 579–605.

Baumol, W. (1988): *Contestable Markets and the Theory of Industry Structure*. Harcourt College Pub.

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.

[45] If such a λ did not exist, one could find i and a sequence $\hat{\mu}_k^{-n} \to \mu^{-n}$ such that $\liminf_k \hat{\mu}_k^{-n}\left(W_i^{-n}\right) \leq \left(1 - \varepsilon\right)\mu^{-n}\left(W_i^{-n}\right)$. But, by the Portmanteau Theorem (Billingsley (2013), Theorem 2.1, part *(iv)*), since W_i^{-n} is open, $\liminf_k \hat{\mu}_k^{-n}\left(W_i^{-n}\right) \geq \mu^{-n}\left(W_i^{-n}\right)$, a contradiction.

Billingsley, P. (2013): *Convergence of Probability Measures*. John Wiley & Sons.

Chamberland, M. (2015): *Single Digits: In Praise of Small Numbers*. Princeton University Press.

Champsaur, P., and J.-C. Rochet (1989): "Multiproduct Duopolists," *Econometrica*, pp. 533–557.

Choi, M., and L. Smith (2017): "Ordinal Aggregation Results via Karlin's Variation Diminishing Property," *Journal of Economic theory*, 168, 1–11.

Crawford, V. P., and E. M. Knoer (1981): "Job Matching with Heterogeneous Firms and Workers," *Econometrica*, pp. 437–450.

Dupuit, J. (1962): "On Tolls and Transport Charges," *International Economic Papers*, 11.

Eeckhout, J., and P. Kircher (2018): "Assortative Matching with Large Firms," *Econometrica*, 86, 85–132.

Epstein, L., and M. Peters (1999): "A Revelation Principle for Competing Mechanisms," *Journal of Economic theory*, 88, 119–160.

Guerrieri, V., R. Shimer, and R. Wright (2010): "Adverse Selection in Competitive Search Equilibrium," *Econometrica*, 78(6), 1823–1862.

Jullien, B. (2000): "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93(1), 1–47.

Kelso, A. S., and V. P. Crawford (1982): "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, pp. 1483–1504.

Laffont, J.-J., and D. Martimort (2002): *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Lester, B., A. Shourideh, V. Venkateswaran, and A. Zetlin-Jones (2018): "Screening and Adverse Selection in Frictional Markets," *Journal of Political Economy*, forthcoming.

Liu, Q., G. J. Mailath, A. Postlewaite, and L. Samuelson (2014): "Stable Matching with Incomplete Information," *Econometrica*, 82(2), 541–587.

Maggi, G., and A. Rodriguez-Clare (1995): "On Countervailing Incentives," *Journal of Economic Theory*, 66(1), 238–263.

Mardsen, J., and A. Tromba (2012): *Vector Calculus*. W. H. Freeman and Company, New York, sixth edition edn.

Martimort, D., and L. Stole (2002): “The Revelation and Delegation Principles in Common Agency Games,” *Econometrica*, 70, 1659–1673.

Maskin, E., and J. Riley (1984): “Monopoly with Incomplete Information,” *The RAND Journal of Economics*, 15(2), 171–196.

Mussa, M., and S. Rosen (1978): “Monopoly and Product Quality,” *Journal of Economic theory*, 18(2), 301–317.

Quah, J., and B. Strulovici (2012): “Aggregating the Single Crossing Property,” *Econometrica*, 80, 2333–2348.

Reny, P. (1999): “On the Existence of Pure and Mixed Strategy Nash Equilibria in Discontinuous Games,” *Econometrica*, 67, 1029–1056.

Rochet, J.-C., and L. Stole (2002): “Nonlinear Pricing with Random Participation,” *The Review of Economic Studies*, 69, 277–311.

Rothschild, M., and J. Stiglitz (1976): “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information,” *The Quarterly Journal of Economics*, pp. 629–649.

Rudin, W. (1987): *Real and Complex Analysis*. McGraw Hill, third edn.

Salop, S. (1979): “Monopolistic Competition with Outside Goods,” *The Bell Journal of Economics*, pp. 141–156.

Spulber, D. F. (1989): “Product Variety and Competitive Discounts,” *Journal of Economic Theory*, 48(2), 510–525.

Stole, L. A. (1995): “Nonlinear Pricing and Oligopoly,” *Journal of Economics & Management Strategy*, 4(4), 529–562.

——— (2007): “Price Discrimination and Competition,” *Handbook of industrial organization*, 3, 2221–2299.

Online Appendix for "Screening in Vertical Oligopolies"

Hector Chade and Jeroen Swinkels

1 Omitted Proofs for Section 4.1

1.1 Proofs for Section 4.1.1

Proposition 5 *Fix n, s^{-n}, and $s^n = (\alpha, v)$. Let $P \equiv \{\theta | \pi(\theta, \alpha, v) \geq 0\}$. Then, there is $(\hat{\alpha}, \hat{v})$ with $\pi(\cdot, \hat{\alpha}, \hat{v}) \geq 0$ that agrees on P with (α, v). If (α, v) is a best response to s^{-n}, then $\pi(\theta, \alpha, v) \geq 0$ for almost all θ where $\varphi > 0$.*

Proof The idea is simply to remove all menu items for which θ is not in P. Let us first show that P can be taken to be closed. Fix n, and let $G(\theta, v)$ be the subdifferential to v at θ. Since v is convex, G is singleton-valued almost everywhere, and every selection from G is increasing. Thus, since G is compact-valued, it is wlog to assume that $\alpha(\theta) \in \arg\max_{a \in G(\theta, v)} \pi(\theta, a, v)$ for all θ. But then, since G is upper hemicontinuous in θ, $\pi(\cdot, \alpha, v)$ is upper semicontinuous (Aliprantis and Border (2006), Lemma 17.30, p. 569), and so P is a closed subset of $[0, 1]$, and hence compact.

Now let's build the menu that results when menu items with θ not in P are removed. For each $\theta' \in [0, 1]$, let $v_L(\cdot, \theta')$ be the line given by $v_L(\theta, \theta') = v(\theta') + (\theta - \theta')\alpha(\theta')$ for all $\theta \in [0, 1]$. Note that $v_L(\theta, \theta) = v(\theta)$, and that since v is convex with $\alpha(\theta') \in G(\theta', v)$, $v_L(\cdot, \theta')$ lies below v for each θ', and that along $v_L(\cdot, \theta')$, the profits to the firm are constant using private values and since the action is constant. If P is empty, set $(\alpha, v) = (\alpha_*^n, v_*^n)$, and we are done. If P is non-empty, define $\hat{v}(\theta) = \max_{\theta' \in P} v_L(\theta, \theta')$. Then, $\hat{v}$, which is the maximum of a set of lines, is convex, with $\hat{v} \leq v$ (since each $v_L(\cdot, \theta')$ lies below v) and $\hat{v} = v$ on P (using that $v_L(\theta, \theta) = v(\theta)$). Let $\hat{\alpha}$ be a selection from $G(\cdot, \hat{v})$, where we can take $\hat{\alpha} = \alpha$ on P, and where at any $\theta \notin P$, we can take $\hat{\alpha}(\theta) = \alpha(\theta')$ for some $\theta' \in \arg\max_{\theta' \in P} v_L(\theta, \theta')$. Then by using $(\hat{\alpha}, \hat{v})$, the firm implements the same action on P at the same profit as before (the types in P have no new deviations available), and the firm earns positive profits on any other type, since that type either leaves or, if served, is now imitating a type in P.

Note finally that if (α, v) is a best response to s^{-n}, and $\pi(\theta, \alpha, v) < 0$ for some positive measure set of θ where $\varphi > 0$, then $(\hat{\alpha}, \hat{v})$ gives strictly higher profits than (α, v), a contradiction. $\square$

1.2 Proofs for Section 4.1.4

Details for the Proof of Lemma 4 Let $\hat{s}(q, \varepsilon) = (\alpha(\cdot, q, \varepsilon), v(\cdot, q, \varepsilon))$ be the menu described in Appendix A, and note that for $\theta \in [\theta_J - \varepsilon, \theta_J)$, $\alpha_q(\theta, q, \varepsilon) = 1$, and $v_q(\theta, q, \varepsilon) \leq \varepsilon$. Hence,

$$\frac{\partial}{\partial q}\pi(\theta, s(q, \varepsilon)) \geq \pi_a(\theta, s(q, \varepsilon)) - \varepsilon \geq \pi_a(\theta, \underline{a} + q, v(q)) - \varepsilon,$$

since π is concave in a. Similarly, for $[\theta_J, \theta_J + \varepsilon]$

$$\frac{\partial}{\partial q}\pi\left(\theta, s\left(q,\varepsilon\right)\right) \geq -\pi_a\left(\theta, s\left(q,\varepsilon\right)\right) - \varepsilon \geq -\pi_a\left(\theta, \bar{a} - q, v\left(q\right)\right) - \varepsilon.$$

Hence, recalling that $=_s$ means "has strictly the same sign as,"

$$\begin{aligned}
\frac{\partial}{\partial q}\Pi\left(\hat{s}\left(q,\varepsilon\right), s^{-n}\right) &\geq \int_{\theta_J-\varepsilon}^{\theta_J}\left(\pi_a\left(\theta, \underline{a}+q, v\left(q\right)\right)-\varepsilon\right)h\left(\theta\right)d\theta + \int_{\theta_J}^{\theta_J+\varepsilon}\left(-\pi_a\left(\theta, \bar{a}-q, v\left(q\right)\right)-\varepsilon\right)h\left(\theta\right)d\theta \\
&= \varepsilon\left[\left(\pi_a\left(\theta', \underline{a}+q, v\left(q\right)\right)-\varepsilon\right)h\left(\theta'\right) - \left(\pi_a\left(\theta'', \bar{a}-q, v\left(q\right)\right)-\varepsilon\right)h\left(\theta''\right)\right] \\
&\underset{s}{=} \left(\pi_a\left(\theta', \underline{a}+q, v\left(q\right)\right)-\varepsilon\right)h\left(\theta'\right) - \left(\pi_a\left(\theta'', \bar{a}-q, v\left(q\right)\right)-\varepsilon\right)h\left(\theta''\right) \\
&\cong \left(\pi_a\left(\theta_J, \underline{a}+q, v\left(q\right)\right) - \pi_a\left(\theta_J, \bar{a}-q, v\left(q\right)\right)\right)h\left(\theta_J\right) \\
&> 0
\end{aligned}$$

where the first equality uses the Mean Value Theorem for some $\theta' \in [\theta_J - \varepsilon, \theta_J]$ and $\theta'' \in [\theta_J, \theta_J + \varepsilon]$, where the approximation is arbitrarily good when ε is small, and where the last inequality holds for $q < (\bar{a} - \underline{a})/2$. But then, for ε and q small, $\partial\Pi\left(\hat{s}\left(q,\varepsilon\right), s^{-n}\right)/\partial q > 0$, and we are done. □

1.3 Proofs for Section 4.1.5

Let us first re-express the profits of the firm in a useful and standard way.

Lemma 23 *Fix n, and for any feasible α and v, define*

$$M(\theta, \alpha, v) = V(\alpha(\theta)) + \alpha\left(\theta\right)\theta - v(\theta_l) - \alpha(\theta)\frac{H(\theta_h) - H(\theta)}{h(\theta)}. \tag{18}$$

Then, $\int_{\theta_l}^{\theta_h}\pi(\theta, \alpha, v)h(\theta)d\theta = \int_{\theta_l}^{\theta_h}M(\theta, \alpha, v)h(\theta)d\theta$.

Proof For any α and v, $\int_{\theta_l}^{\theta_h}\pi(\theta, \alpha, v)h(\theta)d\theta = \int_{\theta_l}^{\theta_h}\left(V(\alpha(\theta)) + \alpha\left(\theta\right)\theta - v(\theta_l) - \int_{\theta_l}^{\theta}\alpha(\tau)d\tau\right)h(\theta)d\theta$, and, integrating by parts, $\int_{\theta_l}^{\theta_h}\left(\int_{\theta_l}^{\theta}\alpha\left(\tau\right)d\tau\right)h(\theta)d\theta = \int_{\theta_l}^{\theta_h}\alpha(\theta)(H(\theta_h) - H(\theta))d\theta$. Substituting and rearranging yields the result. □

Proof of Lemma 1 Existence is standard and uniqueness follows since the set of feasible strategies is convex, and the objective function is strictly concave (since $\pi\left(\theta, a, v\right)$ is strictly concave in a and linear in v for each θ). Fix (θ_l, θ_h), and fix the optimum $\tilde{s} = (\tilde{\alpha}, \tilde{v})$.

Step 1 Let us show that there is η such that for all $\theta \in (\theta_l, \theta_h)$, $\pi_a\left(\theta, \tilde{s}\right) = (-\eta + H\left(\theta_h\right) - H\left(\theta\right))/h\left(\theta\right)$,where we can then take $\kappa = -\eta + H\left(\theta_h\right)$. To see the idea, choose any point θ in (θ_l, θ_h).

We will consider perturbations which subtract a small amount from the action schedule near θ, and replace it just to the left of θ_h. We can do this without worrying about monotonicity since this is the relaxed problem. The perturbation has cost $\pi_a(\theta, \tilde{s})\, h(\theta)$ near θ, benefit $\pi_a(\theta_h, \tilde{s})\, h(\theta_h)$ near θ_h, and benefit $H(\theta_h) - H(\theta)$ because v is lowered between θ and θ_h. Setting the cost equal to the benefit, we have $-\pi_a(\theta, \tilde{s})\, h(\theta) + H(\theta_h) - H(\theta) + \pi_a(\theta_h, \tilde{s})\, h(\theta_h) = 0$, and so

$$\pi_a(\theta, \tilde{s}) = \frac{H(\theta_h) + \pi_a(\theta_h, \tilde{s})\, h(\theta_h) - H(\theta)}{h(\theta)} = \frac{-\eta + H(\theta_h) - H(\theta)}{h(\theta)},$$

where $\eta = -\pi_a(\theta_h, \tilde{s})\, h(\theta_h)$.

To formalize this fix $\theta' \in (\theta_l, \theta_h)$ and $0 < \varepsilon < \min\{\theta_h - \theta', \theta' - \theta_l\}/2$. Define $\hat{\alpha}(\cdot, y, \varepsilon)$ to be $\tilde{\alpha} - y/\varepsilon$ on $[\theta' - \varepsilon, \theta']$, $\tilde{\alpha} + y/\varepsilon$ on $[\theta_h - \varepsilon, \theta_h]$, and $\tilde{\alpha}$ elsewhere, and define $\hat{v}(\theta, y, \varepsilon) = \tilde{v}(\theta_l) + \int_{\theta_l}^{\theta} \hat{\alpha}(\tau, y, \varepsilon) d\tau$, noting that $\hat{v}(\theta_h, y, \varepsilon) = \tilde{v}(\theta_h)$, and so for each y, $\hat{s}(y, \varepsilon) = (\hat{\alpha}(\cdot, y, \varepsilon), \hat{v}(\cdot, y, \varepsilon))$ is feasible in $\mathcal{P}(\theta_l, \theta_h)$. Note that $\hat{v}_y(\theta, y, \varepsilon) = -1$ on $[\theta', \theta_h - \varepsilon]$, and $\hat{v}_y(\theta, y, \varepsilon) \in [-1, 0]$ on $[\theta' - \varepsilon, \theta']$ and $[\theta_h - \varepsilon, \theta_h]$.

Let the profit of this perturbation be $j(y, \varepsilon) = \int_{\theta_l}^{\theta_h} \pi(\theta, \hat{s}(y, \varepsilon)) h(\theta) d\theta$. Then, since $\pi_v = -1$,

$$\begin{aligned} j_y(y, \varepsilon) = & \int_{\theta' - \varepsilon,}^{\theta'} (-\pi_a(\theta, \hat{s}(y, \varepsilon))\frac{1}{\varepsilon} - \hat{v}_y(\theta, y, \varepsilon)) h(\theta) d\theta + H(\theta_h - \varepsilon) - H(\theta') \\ & + \int_{\theta_h - \varepsilon,}^{\theta_h} (\pi_a(\theta, \hat{s}(y, \varepsilon))\frac{1}{\varepsilon} - \hat{v}_y(\theta, y, \varepsilon)) h(\theta) d\theta, \end{aligned}$$

where between θ' and $\theta_h - \varepsilon$ we use $\hat{\alpha}_y = 0$ and $\hat{v}_y = -1$. Note that $\hat{s}(0, \varepsilon) = (\tilde{\alpha}, \tilde{v})$. Hence, evaluating $j_y(y, \varepsilon)$ at $y = 0$, and using the Mean Value Theorem, there is $\tau' \in [\theta' - \varepsilon, \theta']$ and $\tau_h \in [\theta_h - \varepsilon, \theta_h]$ such that

$$\begin{aligned} j_y(0, \varepsilon) = & \,\varepsilon \left(-\pi_a(\tau', \tilde{\alpha}, \tilde{v})\frac{1}{\varepsilon} - \hat{v}_y(\tau', 0, \varepsilon) \right) h(\tau') + \left(H(\theta_h - \varepsilon) - H(\theta') \right) \\ & + \varepsilon \left(\pi_a(\tau_h, \tilde{\alpha}, \tilde{v})\frac{1}{\varepsilon} - \hat{v}_y(\tau_h, 0, \varepsilon) \right) h(\tau_h). \end{aligned}$$

But then, since $\hat{v}_y(\tau', 0, \varepsilon)$ and $\hat{v}_y(\tau_h, 0, \varepsilon)$ are bounded,

$$\lim_{\varepsilon \to 0} j_y(0, \varepsilon) = -\pi_a(\theta', \tilde{\alpha}, \tilde{v}) h(\theta') + H(\theta_h) - H(\theta') + \pi_a(\theta_h, \tilde{\alpha}, \tilde{v}) h(\theta_h) = 0,$$

since the perturbation is feasible for y in a neighborhood of zero. Rearranging and taking $\eta = -\pi_a(\theta_h, \tilde{\alpha}, \tilde{v}) h(\theta_h)$, we are done.

Step 2 Let us next show that if one fixes surplus to equal $\tilde{v}(\theta_l)$ at θ_l, and then varies κ, ignoring (3), then profits are single-peaked at $\kappa = H(\theta_h)$. Similarly, if one fixes surplus to equal $\tilde{v}(\theta_h)$ at

θ_h, and then varies κ, ignoring (2), then profits are single-peaked at $\kappa = H(\theta_l)$.

To formalize this, let $v_l(\theta,\kappa) = \tilde{v}(\theta_l) + \int_{\theta_l}^{\theta} \gamma(\tau,\kappa)d\tau$, and let $s_l(\kappa) = (\gamma(\cdot,\kappa), v_l(\cdot,\kappa))$. Since $v_l(\theta_l,\kappa) = \tilde{v}(\theta_l)$ and so is independent of κ, it follows from (18) that on (θ_l,θ_h),

$$\frac{d}{d\kappa}M(\theta, s_l(\kappa)) = \left(\pi_a(\theta, s_l(\kappa)) - \frac{H(\theta_h)-H(\theta)}{h(\theta)}\right)\gamma_\kappa(\theta,\kappa) = \left(\frac{\kappa - H(\theta_h)}{h(\theta)}\right)\gamma_\kappa(\theta,\kappa) =_s -(\kappa - H(\theta_h)),$$

where the equality follows by Step 1, and the equality in sign since $\gamma_\kappa < 0$. But then, letting $Y_l(\kappa) \equiv \int_{\theta_l}^{\theta_h} \pi(\theta, s_l(\kappa))h(\theta)d\theta$, by Lemma 23, $dY_l(\kappa)/d\kappa =_s -(\kappa - H(\theta_h))$, and so $Y_l(\kappa)$ is strictly single-peaked at $\kappa = H(\theta_h)$.

Similarly, if we define $v_h(\theta,\kappa) = \tilde{v}(\theta_h) - \int_{\theta}^{\theta_h} \gamma(\tau,\kappa)d\tau$, then $Y_h(\kappa) \equiv \int_{\theta_l}^{\theta_h} \pi(\theta, \gamma(\cdot,\kappa), v_h(\cdot,\kappa))h(\theta)d\theta$ is strictly single-peaked in κ with maximum at $\kappa = H(\theta_l)$ where to show this, one integrates

$$\int_{\theta_l}^{\theta_h} \pi(\theta,\alpha,v)h(\theta)d\theta = \int_{\theta_l}^{\theta_h}\left(V(\alpha(\theta)) + \alpha(\theta)\theta - v(\theta_h) + \int_{\theta}^{\theta_h}\alpha(\tau)d\tau\right)h(\theta)d\theta$$

by parts to arrive at an analogue to M.

Step 3 Finally, let us show that $\kappa_o = \tilde{\kappa}(\theta_l,\theta_h)$. Note that one of (2) and (3) must bind, otherwise reducing $\tilde{v}$ by a small positive constant, holding fixed $\tilde{\alpha}$, is profitable. Assume that $\tilde{v}(\theta_h) > v^{-n}(\theta_h)$. Then, $s_l(\kappa)$ is feasible for κ on a neighborhood of κ_o, and so, since by Step 2, Y_l is strictly single-peaked with maximum at $H(\theta_h)$ we have $\kappa_o = H(\theta_h)$. Let us see that $\tilde{\kappa}(\theta_l,\theta_h) = H(\theta_h)$ as well, so that $\kappa_o = \tilde{\kappa}(\theta_l,\theta_h)$. Since $\tilde{v}(\theta_l) = v^{-n}(\theta_l)$, and since $(\tilde{\alpha},\tilde{v}) = (\gamma(\cdot,\kappa_o),\tilde{v})$ is feasible, we have $\tilde{v}(\theta_h) = v^{-n}(\theta_l) + \int_{\theta_l}^{\theta_h}\gamma(\tau, H(\theta_h))d\tau > v^{-n}(\theta_h)$, and so $\iota(\theta_l,\theta_h,H(\theta_h)) < 0$, and thus by definition of $\tilde{\kappa}$, we have $\tilde{\kappa}(\theta_l,\theta_h) = H(\theta_h)$ as well. Similarly, if $\tilde{v}(\theta_l) > v^{-n}(\theta_l)$ then, using Y_h, we must have $\kappa_o = H(\theta_l) = \tilde{\kappa}(\theta_l,\theta_h)$.

Assume finally that (2) and (3) both bind. Then, by definition, $\iota(\theta_l,\theta_h,\kappa_o) = 0$. Assume $\kappa_o > H(\theta_h)$. Then,

$$v_l(\theta_h, H(\theta_h)) = \tilde{v}(\theta_l) + \int_{\theta_l}^{\theta_h}\gamma(\tau, H(\theta_h))d\tau > \tilde{v}(\theta_l) + \int_{\theta_l}^{\theta}\gamma(\tau,\kappa_o)d\tau = \tilde{v}(\theta_h) = v^{-n}(\theta_h),$$

so that $s_l(H(\theta_h))$ is feasible, which contradicts the optimality of $(\tilde{\alpha},\tilde{v})$ since Y_l is uniquely maximized at $H(\theta_h)$, and Y_l ignores (3). Hence $\kappa_o \leq H(\theta_h)$. Similarly, $\kappa_o \geq H(\theta_l)$, and thus $\kappa_o \in [H(\theta_l), H(\theta_h)]$, from which $\kappa_o = \tilde{\kappa}(\theta_l,\theta_h)$, again by definition of $\tilde{\kappa}$. □

We now prove that any optimum of the original problem has the form given by Lemma 1.

Proposition 6 *Let s be Nash with NEO. Then, for each n, there is $\kappa^n \in [H(\theta_l^n), H(\theta_h^n)]$ such that $\alpha^n = \gamma^n(\cdot,\kappa^n)$ on (θ_l^n,θ_h^n), where $\kappa^1 = 0$, and $\kappa^N = 1$.*

Proof We will show that if on (θ_l, θ_h), (α, v) is not equal to $(\tilde{\alpha}, \tilde{v})$, the optimal solution to the relaxed problem, then we can profitably perturb (α, v) in the direction of $(\tilde{\alpha}, \tilde{v})$.[46] We need this perturbation to respect monotonicity and the fact that workers both within and outside of (θ_l, θ_h) may be affected. This proof would be substantially simpler if all crossings where transversal, but we know this will fail when firms are not very differentiated.

Let $\check{s}(\delta)$ be given by $\check{\alpha}(\cdot, \delta) = (1-\delta)\alpha + \delta\tilde{\alpha}$ and $\check{v}(\cdot, \delta) = (1-\delta)v + \delta\tilde{v}$, so that $\check{s}(0) = (\alpha, v)$ and $\check{s}(1) = (\tilde{\alpha}, \tilde{v})$. The problem with $\check{s}$ is that when crossings are not transversal, $\check{s}(\delta)$ need not serve all of (θ_l, θ_h) even for small δ. So, let $\bar{v} = v^{-n}/2 + v/2$, so that $\bar{v} > v^{-n}$ on (θ_l, θ_h). Now, let $\grave{v}(\cdot, \delta) = \max(\bar{v}, \check{v}(\cdot, \delta))$, let $\grave{\alpha}(\cdot, \delta)$ be a subgradient to $\grave{v}(\cdot, \delta)$, and let $\grave{s}(\delta) = (\grave{\alpha}(\cdot, \delta), \grave{v}(\cdot, \delta))$. By construction, $\grave{s}$ always wins on (θ_l, θ_h), and may serve other types as well. Also, since on (θ_l, θ_h), $v > v^{-n}$, $\grave{s}(0) = (\alpha, v)$. Finally, let $P(\delta)$ be the set upon which $\grave{s}(\delta)$ is profitable, and construct $\hat{s}(\delta) = (\hat{\alpha}(\cdot, \delta), \hat{v}(\cdot, \delta))$ from $\grave{s}(\delta)$ as in Proposition 5. We then have

$$\begin{aligned}\Pi(\hat{s}(\delta), s^{-n}) &= \int \pi(\theta, \hat{s}(\delta))\varphi(\theta, \hat{s}(\delta), s^{-n})h(\theta)d\theta \geq \int_{P(\delta)\cap(\theta_l,\theta_h)} \pi(\theta, \hat{s}(\delta))\varphi(\theta, \hat{s}(\delta), s^{-n})h(\theta)d\theta \\ &= \int_{P(\delta)\cap(\theta_l,\theta_h)} \pi(\theta, \grave{s}(\delta))h(\theta)d\theta \geq \int_{\theta_l}^{\theta_h} \pi(\theta, \grave{s}(\delta))h(\theta)d\theta.\end{aligned}$$

The first inequality follows since $\pi(\cdot, \hat{s}(\delta)) \geq 0$, the second equality since $\hat{s}(\delta)$ and $\grave{s}(\delta)$ agree on $P(\delta)$ and $\varphi(\cdot, \grave{s}(\delta)) = 1$ on (θ_l, θ_h), and the second inequality since $\pi(\theta, \grave{s}(\delta)) \leq 0$ outside of $P(\delta)$.

It is thus enough to show that for δ sufficiently small, $\int_{\theta_l}^{\theta_h} \pi(\theta, \grave{s}(\delta))h(\theta)d\theta > \int_{\theta_l}^{\theta_h} \pi(\theta, \alpha, v)h(\theta)d\theta$, since by *PS*, $\varphi(\theta, s) = 0$ outside of $[\theta_l, \theta_h]$. Because $\grave{s}(0) = (\alpha, v)$, it is sufficient that $\frac{d}{d\delta}\int_{\theta_l}^{\theta_h} \pi(\theta, \grave{s}(\delta))h(\theta)d\theta\Big|_{\delta=0} > 0$. But, $\frac{d}{d\delta}\int_{\theta_l}^{\theta_h} \pi(\theta, \grave{s}(\delta))h(\theta)d\theta\Big|_{\delta=0} = \frac{d}{d\delta}\int_{\theta_l}^{\theta_h} \pi(\theta, \check{s}(\delta))h(\theta)d\theta\Big|_{\delta=0}$, since for each $\theta \in (\theta_l, \theta_h)$, $v(\theta) > \bar{v}(\theta)$, and so at $\delta = 0$, $(\check{\alpha}(\theta, \delta))_\delta = (\grave{\alpha}(\theta, \delta))_\delta$ and $(\check{v}(\theta, \delta))_\delta = (\grave{v}(\theta, \delta))_\delta$. And, since $(\tilde{\alpha}, \tilde{v})$ is the unique solution on (θ_l, θ_h) to the relaxed problem $\mathcal{P}(\theta_l, \theta_h)$, and since $\check{s}(0) = (\alpha, v)$, and so is feasible in $\mathcal{P}(\theta_l, \theta_h)$,

$$\int_{\theta_l}^{\theta_h} \pi(\theta, \tilde{\alpha}, \tilde{v})h(\theta)d\theta = \int_{\theta_l}^{\theta_h} \pi(\theta, \check{s}(1))h(\theta)d\theta > \int_{\theta_l}^{\theta_h} \pi(\theta, \check{s}(0))h(\theta)d\theta = \int_{\theta_l}^{\theta_h} \pi(\theta, \alpha, v)h(\theta)d\theta.$$

Now, $\check{s}$ is linear in δ, and $\pi(\theta, \cdot, \cdot)$ is concave in the action and utility, and thus $\int_{\theta_l}^{\theta_h} \pi(\theta, \check{s}(\delta))h(\theta)d\theta$ is concave in δ. But then, by the previous strict inequality, it must be that $\frac{d}{d\delta}\int_{\theta_l}^{\theta_h} \pi(\theta, \check{s}(\delta))h(\theta)d\theta\Big|_{\delta=0} > 0$.

Finally, let us see that $\kappa^N = 1$ (where the proof that $\kappa^1 = 0$ is similar). Note first that for $\theta \geq \theta_l^N$, $v^{-N} = v^{N-1}$. Thus, by definition $v^N(\theta_l^N) = v^{N-1}(\theta_l^N)$. But, by *NEO*, for all $\theta > \theta_l^N$, we have $\alpha^N(\theta) > \alpha^N(\theta_l^N) \geq a_e^{N-1} \geq \alpha^{N-1}(\theta)$ and thus $v_\theta^N > v_\theta^{N-1}$. Thus, $v^N(1) > v^{N-1}(1)$,

[46] It is not important how $(\tilde{\alpha}, \tilde{v})$ is defined outside of (θ_l, θ_h) so long as monotonicity, continuity of actions, and the integral condition hold.

and hence $\iota\left(\theta_l^N, 1, \kappa^N\right) < 0$, which by definition of $\tilde{\kappa}$ can only hold if $\kappa^N = H(1) = 1$. $\square$

1.4 Proofs for Section 4.1.6

Proposition 7 *Let s be Nash with NEO. Then,* (5) *holds.*

Proof Fix n. We will prove (5) for θ_h, with the case at θ_l analogous. We will consider perturbations that add or subtract workers in a continuous fashion immediately to the right or left of θ_h. We need to respect monotonicity and the integral condition, and ensure that our perturbed menus continue to serve an interval of workers (as opposed to a disconnected set thereof).[47]

If $\theta_l^{n+1} > \theta_h^n = \theta_h$, then (5) is automatic, since by Corollary 1 and the definition of PS, $\pi(\theta_h, \alpha, v) = 0$ and $\alpha(\theta_h) = \alpha^{n+1}(\theta_h)$. So, assume $\theta_l^{n+1} = \theta_h$, and note that by Proposition 6, α is strictly increasing to the left of θ_h, and $a^{-n} = \alpha^{n+1}$ is strictly increasing to the right of θ_h.

Step 1 Let us first define a basic perturbation $(\hat{\alpha}(\cdot, y), \hat{v}(\cdot, y))$ indexed by y. Fix n and $0 < \varepsilon < \theta_h - \theta_l$. For y positive or negative, define $\hat{\alpha}(\theta, y)$ as $\alpha(\theta)$ if $\theta < \theta_h - \varepsilon$, and $\max\{\alpha(\theta_h - \varepsilon), \min\{\alpha(\theta) + y, \alpha(\theta_h)\}\}$ if $\theta \geq \theta_h - \varepsilon$. That is, above $\theta_h - \varepsilon$, change actions by y but censor them to be above $\alpha(\theta_h - \varepsilon)$ and below $\alpha(\theta_h)$. Note that $\hat{\alpha}$ is continuous and $\hat{\alpha}(\cdot, y)$ is increasing. Define $\hat{v}(\theta, y) = v(\theta_l) + \int_{\theta_l}^{\theta} \hat{\alpha}(\tau, y) d\tau$. Because $\hat{\alpha}(\tau, y)$ is bounded and for each y, differentiable in y for almost all τ, with $\hat{\alpha}_y(\tau, y) \in \{0, 1\}$ wherever it is defined, $\hat{v}$ is continuously differentiable in (θ, y) wherever $\theta > \theta_h - \varepsilon$, with $\hat{v}_y(\theta_h, 0) = \varepsilon > 0$.

Step 2 Let us now use the basic perturbation to add or subtract types near θ_h. Define $\hat{y}(\theta')$ implicitly by $\hat{v}(\theta', \hat{y}(\theta')) - v^{-n}(\theta') = 0$. Then $\hat{y}$ is well defined on an interval around θ_h, with

$$\hat{y}_{\theta'}(\theta') = \frac{a^{-n}(\theta') - \hat{\alpha}(\theta', \hat{y}(\theta'))}{\hat{v}_y(\theta', \hat{y}(\theta'))} \geq 0. \tag{19}$$

Further, when $\hat{y}(\theta') > 0$, then $\hat{v}(\theta, \hat{y}(\theta')) - v^{-n}(\theta) > 0$ for all $\theta \in (\theta_l, \theta_h]$, and hence any crossing of zero by $\hat{v}(\cdot, \hat{y}(\theta')) - v^{-n}(\cdot)$ above θ_l occurs where $\theta > \theta_h$, and thus where

$$(\hat{v}(\theta, \hat{y}(\theta')) - v^{-n}(\theta))_\theta = \hat{\alpha}(\theta, \hat{y}(\theta')) - a^{-n}(\theta) = \alpha(\theta_h) - a^{-n}(\theta) < 0,$$

since $a^{-n}(\theta) > a^{-n}(\theta_h) \geq \alpha(\theta_h)$. Thus, indeed θ' is the unique crossing, and so $\varphi = 1$ for all $\theta \in (\theta_l, \theta')$, and $\varphi = 0$ outside of $[\theta_l, \theta']$. Similarly, if $\hat{y}(\theta') < 0$, then any crossing of zero by $\hat{v}(\cdot, \hat{y}(\theta')) - v^{-n}(\cdot)$ above θ_l occurs where $\theta \in (\theta_h - \varepsilon, \theta_h)$, and thus where $\hat{\alpha}(\theta, \hat{y}(\theta')) < \alpha(\theta)$, and hence

$$(\hat{v}(\theta, \hat{y}(\theta')) - v^{-n}(\theta))_\theta = \hat{\alpha}(\theta, \hat{y}(\theta')) - a^{-n}(\theta) < \alpha(\theta) - a^{-n}(\theta) < 0,$$

[47] This proof would be much easier if all crossing were strictly transversal. Then, we could use $\gamma(\cdot, \kappa)$ and vary κ holding fixed $v(\theta_l)$.

since $\alpha(\theta) < \alpha(\theta_h) \leq a_e^n \leq a^{-n}(\theta)$ by *NEO*, and so again $\varphi = 1$ for all $\theta \in (\theta_l, \theta')$, and $\varphi = 0$ outside of $[\theta_l, \theta']$.

Step 3 Since this perturbation is feasible, it must be unprofitable. Let us show that this implies (5). To do so, let $j(\theta')$ be the profit from the perturbation. Then,

$$j(\theta') = \int_{\theta_l}^{\theta_h - \varepsilon} \pi(\theta, \alpha, v) h(\theta) d\theta + \int_{\theta_h - \varepsilon}^{\theta'} \pi(\theta, \hat{\alpha}(\cdot, \hat{y}(\theta')), \hat{v}(\cdot, \hat{y}(\theta')) h(\theta) d\theta,$$

since for $\theta < \theta_h - \varepsilon$, $\hat{\alpha} = \alpha$ and $\hat{v} = v$. Thus,

$$\begin{aligned} j_{\theta'}(\theta') &= \pi(\theta', \hat{\alpha}(\cdot, \hat{y}(\theta')), \hat{v}(\cdot, \hat{y}(\theta')) h(\theta') \\ &\quad + \hat{y}_\theta(\theta') \int_{\theta_h - \varepsilon}^{\theta'} \left(\pi_a(\theta, \hat{\alpha}(\cdot, \hat{y}(\theta')), \hat{v}(\cdot, \hat{y}(\theta')) \hat{\alpha}_y(\theta, \hat{y}(\theta')) - \hat{v}_y(\theta, \hat{y}(\theta'))\right) h(\theta) d\theta. \end{aligned}$$

To evaluate this at $\theta' = \theta_h$, note that $\hat{y}(\theta_h) = 0$, $\hat{\alpha}(\theta, 0) = \alpha(\theta)$, $\hat{\alpha}_y^n(\theta, 0) = 1$ for $\theta \in (\theta_h - \varepsilon, \theta_h)$, and zero outside of $[\theta_h - \varepsilon, \theta_h]$ and that $\hat{v}(\cdot, 0) = v$, and so, using (19) and $\hat{v}_y(\theta_h, 0) = \varepsilon$,

$$\begin{aligned} j_{\theta'}(\theta_h) &= \pi(\theta_h, \alpha, v) h(\theta_h) + \frac{a^{-n}(\theta_h) - \alpha(\theta_h)}{\varepsilon} \int_{\theta_h - \varepsilon}^{\theta_h} (\pi_a(\theta, \alpha, v) - \hat{v}_y(\theta, 0)) h(\theta) d\theta \\ &= \pi(\theta_h, \alpha, v) h(\theta_h) + \left(a^{-n}(\theta_h) - \alpha(\theta_h)\right) (\pi_a(\tau, \alpha, v) - \hat{v}_y(\tau, 0)) h(\tau) \end{aligned}$$

for some $\tau \in [\theta_h - \varepsilon, \theta_h]$ by the Mean Value Theorem, and where we note that $\hat{v}_y(\tau, 0) = \tau - (\theta_h - \varepsilon) \in [0, \varepsilon]$. Since (α, v) is optimal, we have $j_{\theta'}(\theta_h) = 0$. Taking $\varepsilon \to 0$, we have $\tau \to \theta_h$, and hence, canceling $h(\theta_h)$, we arrive at $0 = \pi(\theta_h, \alpha, v) + (a^{-n}(\theta_h) - \alpha(\theta_h)) \pi_a(\theta_h, \alpha, v)$. Thus (5) holds, and we are done. □

2 Numerical Analysis and Figure 1

We take four firms, with $\mathcal{V}^n(a) = \zeta^n + \beta^n \log(\rho + a)$, where with some mild abuse we set $\rho = 0$. Assume that $\mathcal{V}(a, \theta) = -(3 - \theta) a$, and that h is uniform on $[0, 1]$. From *IO*, $\gamma^n(\theta, \kappa^n) = \beta^n / (3 + \kappa^n - 2\theta)$ for $n = 1, ..., 4$, and so stacking holds if $\beta^{n+1}/\beta^n > 2$. So assume $\beta^1 = 1, \beta^2 = 4, \beta^3 = 9, \beta^4 = 20, \zeta^1 = 2.5, \zeta^2 = 3, \zeta^3 = -2$, and $\zeta^4 = -23$, where the values of ζ^n are chosen so that each firm is relevant. Integrating γ^n yields $v^n(\theta) = v^n(0) + (\beta^n/2) \log((3 + \kappa^n)/(3 + \kappa^n - 2\theta))$, and so the nine equations that characterize equilibria are

$$\begin{aligned} v^n(\theta^n) - v^{n+1}(\theta^n) &= 0, \quad n = 1,2,3 \\ \pi^n(\theta^n, \gamma^n(\cdot, \kappa^n), v^n) + (\kappa^n - \theta^n)(\gamma^{n+1}(\theta^n) - \gamma^n(\theta^n)) &= 0, \quad n = 1,2,3 \\ \pi^{n+1}(\theta^n, \gamma^{n+1}(\cdot, \kappa^{n+1}), v^{n+1}) + (\kappa^{n+1} - \theta^n)(\gamma^n(\theta^n) - \gamma^{n+1}(\theta^n)) &= 0, \quad n = 1,2,3, \end{aligned}$$

with nine unknowns $\kappa^2, \kappa^3, \theta^1, \theta^2, \theta^3, v^1(0), v^2(0), v^3(0)$, and $v^4(0)$. Solving numerically, inserting the values for $v^n(0)$ and κ^n into v^n and graphing gives us Figure 1.[48]

3 An Extreme Welfare Reversal Example

Examples where all types prefer complete information are easy to build if one brings h into play. To see this, modify the distribution over types so that there is $\tau < 1/2$ weight at each of $\theta = 0$ and $\theta = 1$, and let the remaining weight be uniform. (The point masses could be dispensed with, but make things simple.) When τ is zero, we are back in the uniform case, while as τ grows, there are more consumers that are essentially captive to one firm or the other, muting competition. Then, it is a routine calculation that that stacking holds for $\beta^1 = 1$, $\beta^2 = 5$, and $\tau = .2$. and it can be calculated numerically that if $k_1 = 4.5$ and $k_2 = 4$, then the solution for $\tau = .2$ is given by $\theta^1 = 0.55129$, $v^1(0) = 1.4169$, and $v^2(0) = 0.94957$. As can be seen from Figure 6, all types prefer the complete information case.

4 Omitted Proofs for Section 5

4.1 Proofs for Section 5.2

Proof of Theorem 4 We proceed in a series of steps.

Step 0 Define $\tilde{a}$ by $V_a(\tilde{a}, \bar{z}) = -1 - (1/h(1))$. Define $z^a(a) = \arg\max V(a, \cdot)$ as the optimal technology to implement action a, and $a^\theta(\theta) = \arg\max V(a, z^a(a)) + a\theta$ as the action that, when implemented with technology $z^a(a)$, maximizes the surplus for type θ. Define $z^\theta(\theta) \equiv z^a(a^\theta(\theta))$, so that type θ is best served by a firm with technology $z^\theta(\theta)$ and action $a^\theta(\theta)$. For any given $\kappa \in [0,1]$, let $\gamma(\cdot, \kappa, z)$ solve (1) with technology z. Then, $\tilde{a}$ is an upper bound for $\gamma(\theta, \kappa, z)$ for all (θ, κ, z) with $\kappa \in [0,1]$, and z^a, a^θ, and z^θ are positive, well-defined, continuously differentiable and bounded away from zero and ∞ for all $\theta \in [0,1]$ and $a \in [0, \tilde{a}]$.

That $\tilde{a}$ is a relevant upper bound on $\gamma(\theta, \kappa, z)$ follows from (1). The properties of z^a, a^θ, and z^θ follow from our ambient assumptions and the Implicit Function Theorem. Formally, note that

[48] The solution is $\kappa^2 = 0.53318$, $\kappa^3 = 0.84976$, $\theta^1 = 0.26105$, $\theta^2 = 0.68527$, $\theta^3 = 0.91815$, $v^1(0) = 0.1502$, $v^2(0) = -0.074$, $v^3(0) = -1.0726$, and $v^4(0) = -4.3006$.

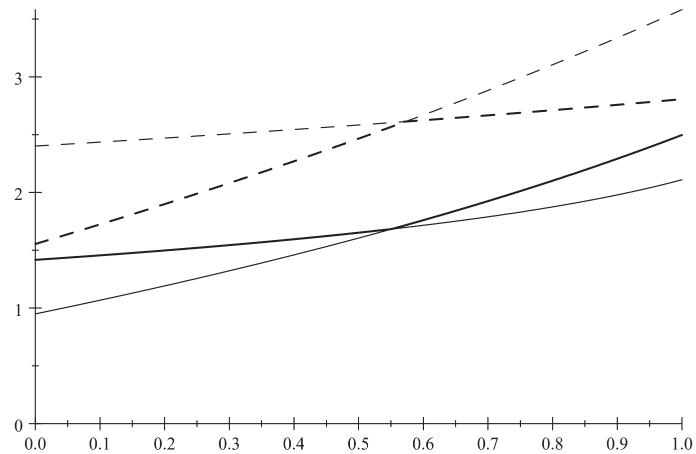

Figure 6: **A Complete Welfare Reversal.** The thin solid lines are the incomplete information equilibrium surplus of the two firms, with the thicker solid line the equilibrium surplus as a function of the type. The thin dashed lines are the efficient surplus each firm can offer, with the thicker dashed line the equilibrium surplus under incomplete information.

for any $\theta \in [0,1]$ and $z \in [0,\bar{z}]$, $\gamma(\theta,\kappa,z)$ uniquely solves $V_a = -\theta + (\kappa - H)/h$. Thus, since $V_{az} > 0$,

$$V_a\left(\gamma\left(\theta,\kappa,z\right),\bar{z}\right) \geq V_a\left(\gamma\left(\theta,\kappa,z\right),z\right) = -\theta + \frac{\kappa - H\left(\theta\right)}{h\left(\theta\right)} \geq -1 - \frac{1}{h\left(1\right)} = V_a(\tilde{a},\bar{z}),$$

and since $V_{aa} < 0$, we obtain that $\gamma(\theta,\kappa,z) \leq \tilde{a}$ for all (θ,z) and $\kappa \in [0,1]$. Note that z^a is implicitly defined by $V_z\left(a,z^a\left(a\right)\right) = 0$ and hence $z^a_a\left(a\right) = -V_{az}\left(a,z^a\left(a\right)\right)/V_{zz}\left(a,z^a\left(a\right)\right) > 0$. Consider $j\left(a,\theta\right) = V\left(a,z^a\left(a\right)\right) + a\theta$. Then, $j_a\left(0,\theta\right) = V_a\left(0,z^a\left(0\right)\right) + \theta \geq V_a\left(0,0\right) > 0$, and for all $a \geq \tilde{a}$, $j_a\left(a,\theta\right) = V_a\left(a,z^a\left(a\right)\right) + \theta \leq V_a\left(a,\bar{z}\right) + 1 \leq V_a\left(\tilde{a},\bar{z}\right) + 1 < 0$. Thus, j has a maximum on $(0,\tilde{a})$. Further,

$$\begin{aligned} j_{aa}\left(a,\theta\right) &= \left(V_a\left(a,z^a\left(a\right)\right)\right)_a = V_{aa}\left(a,z^a\left(a\right)\right) + V_{az}\left(a,z^a\left(a\right)\right)z^a_a\left(a\right) \\ &= V_{aa}\left(a,z^a\left(a\right)\right) - V_{az}\left(a,z^a\left(a\right)\right)\frac{V_{az}\left(a,z^a\left(a\right)\right)}{V_{zz}\left(a,z^a\left(a\right)\right)} \\ &=_s -\left(V_{aa}\left(a,z^a\left(a\right)\right)V_{zz}\left(a,z^a\left(a\right)\right) - V^2_{az}\left(a,z^a\left(a\right)\right)\right) < 0, \end{aligned}$$

since we assumed that the Hessian of V was strictly positive. Hence $j\left(\cdot,\theta\right)$ has a unique maximum $a^\theta\left(\theta\right) \in (0,\tilde{a})$, and the pair $a^\theta\left(\theta\right)$ and $z^\theta\left(\theta\right) \equiv z^a(a^\theta(\theta))$ jointly maximize $V\left(a,z\right) + a\theta$ for each θ, and thus $v_*(\theta) = j(a^\theta(\theta),\theta)$. Differentiating the identity $j_a(a^\theta(\theta),\theta) = 0$ and using the expression for j_{aa}, we have $a^\theta_\theta(\theta) = -1/j_{aa}(a^\theta(\theta),\theta) > 0$, and hence $z^\theta_\theta(\theta) = z^a_a(a^\theta(\theta))a^\theta_\theta(\theta) > 0$.

Since the type space, $[0,1]$, is compact, and since all relevant actions will come from the compact interval $[0,\tilde{a}]$, and technologies from $[0,z]$, and since V is $\mathcal{C}^2$, we have that $\ell_{a^\theta_\theta} > 0$, and $\ell_{z^\theta_\theta} > 0$, where for any function g, we will use ℓ_g as shorthand for the infimum of g over the relevant domain.

Let the maximum surplus a firm with technology z can offer to type θ be $\bar{v}(\theta,z) = \max_a (V(a,z) + \theta a) = V(\bar{a}(\theta,z), z) + \theta\bar{a}(\theta,z)$, where $\bar{a}$ is defined by $V_a(\bar{a}(\theta,z),z) + \theta = 0$, and hence

$$\bar{a}_\theta(\theta,z) = \frac{-1}{V_{aa}(\bar{a}(\theta,z),z)} > 0, \text{ and } \bar{a}_z(\theta,z) = \frac{-V_{az}(\bar{a}(\theta,z),z)}{V_{aa}(\bar{a}(\theta,z),z)} > 0.$$

Each of these is finite on the compact set $[0,1]\times[0,\bar{z}]$, and hence has a strictly positive uniform lower bound, $\ell_{\bar{a}_\theta}$ and $\ell_{\bar{a}_z}$, and finite upper bound, $\upsilon_{\bar{a}_\theta}$ and $\upsilon_{\bar{a}_z}$.

In what follows fix N, $\{z^n\}_{n=1}^N$, and an equilibrium s.

Step 1 Let us show first a lower bound on how much an entering firm can earn as function of how far apart its competitors are. Partition the interval $\left[z^\theta(0), z^\theta(1)\right]$ by those elements of $\{z^n\}_{n=1}^N$ that lie in $\left[z^\theta(0), z^\theta(1)\right]$. Let d_z be the length of the longest element of the partition. We claim there is $\rho_1 \in (0,\infty)$ such that an entrant can earn at least $\rho_1 d_z^3$.

Let $[z_l, z_h]$ be a largest element of the partition, so that $z_h - z_l = d_z$. Associated with $[z_l,z_h]$ is an interval of types $[\theta_l,\theta_h] = [\left(z^\theta\right)^{-1}(z_l), \left(z^\theta\right)^{-1}(z_h)]$ where by the Mean Value Theorem,

$$\theta_h - \theta_l \geq \frac{d_z}{\upsilon_{z^\theta_\theta}}. \tag{20}$$

Let $\tilde{\theta} = (\theta_l + \theta_h)/2$. Since V is concavity and there are no firms in (z_l,z_h), $\max\{\bar{v}(\tilde{\theta},z_l), \bar{v}(\tilde{\theta},z_h)\} \geq v^O(\tilde{\theta})$, since z_l does a better job of serving $\tilde{\theta}$ than any $z < z_l$, and similarly for $z > z_h$, and where, if for example $z_l = z^\theta(0)$, so that z_l need not be an existing firm, then the inequality holds *a fortiori*.

Note that $v_*(\tilde{\theta}) = \bar{v}(\tilde{\theta}, z^\theta(\tilde{\theta}))$, and $\bar{v}(\theta_l, z^\theta(\theta_l)) - \bar{v}(\theta_l, z_l) = 0$, since $z^\theta(\theta_l) = z_l$. Hence

$$\begin{aligned} v_*(\tilde{\theta}) - \bar{v}(\tilde{\theta}, z_l) &= \bar{v}(\tilde{\theta}, z^\theta(\tilde{\theta})) - \bar{v}(\tilde{\theta}, z_l) = \bar{v}(\theta_l, z^\theta(\theta_l)) - \bar{v}(\theta_l, z_l) + \int_{\theta_l}^{\tilde{\theta}} \left(\left(\bar{v}\left(\theta, z^\theta(\theta)\right)\right)_\theta - \bar{v}_\theta(\theta, z_l)\right) d\theta \\ &= \int_{\theta_l}^{\tilde{\theta}} \left(\left(\bar{v}\left(\theta, z^\theta(\theta)\right)\right)_\theta - \bar{v}_\theta(\theta, z_l)\right) d\theta = \int_{\theta_l}^{\tilde{\theta}} \left(\bar{v}_\theta\left(\theta, z^\theta(\theta)\right) - \bar{v}_\theta(\theta, z_l)\right) d\theta \\ &= \int_{\theta_l}^{\tilde{\theta}} \left(\bar{a}\left(\theta, z^\theta(\theta)\right) - \bar{a}(\theta, z_l)\right) d\theta, \end{aligned}$$

where the fourth and fifth equalities use the Envelope Theorem. But, again since $z^\theta(\theta_l) = z_l$,

$\bar{a}(\theta, z^{\theta}(\theta_l)) - \bar{a}(\theta, z_l) = 0$, and so

$$\bar{a}(\theta, z^{\theta}(\theta)) - \bar{a}(\theta, z_l) = \int_{\theta_l}^{\theta} \frac{\partial}{\partial \tau} \bar{a}\left(\theta, z^{\theta}(\tau)\right) d\tau = \int_{\theta_l}^{\theta} \bar{a}_z\left(\theta, z^{\theta}(\tau)\right) z^{\theta}_{\theta}(\tau)\, d\tau \geq (\theta - \theta_l)\, \ell_{\bar{a}_z} \ell_{z^{\theta}_{\theta}}$$

and hence, substituting,

$$v_*(\tilde{\theta}) - \bar{v}(\tilde{\theta}, z_l) \geq \ell_{\bar{a}_z} \ell_{z^{\theta}_{\theta}} \int_{\theta_l}^{\tilde{\theta}} (\theta - \theta_l)\, d\theta = \ell_{\bar{a}_z} \ell_{z^{\theta}_{\theta}} \frac{(\tilde{\theta} - \theta_l)^2}{2} = \frac{\ell_{\bar{a}_z} \ell_{z^{\theta}_{\theta}}}{8} (\theta_h - \theta_l)^2$$

and similarly for z_h. Hence, using (20),

$$v_*(\tilde{\theta}) - \max\{\bar{v}(\tilde{\theta}, z_l), \bar{v}(\tilde{\theta}, z_h)\} \geq \frac{\ell_{\bar{a}_z} \ell_{z^{\theta}_{\theta}}}{8} (\theta_h - \theta_l)^2 \geq \frac{\ell_{\bar{a}_z} \ell_{z^{\theta}_{\theta}}}{8 v^2_{z^{\theta}_{\theta}}} d_z^2 \equiv \delta. \tag{21}$$

Let $\tilde{z} = z(\tilde{\theta})$ enter, offer $\bar{a}(\cdot, \tilde{z})$, and offer surplus $\bar{v}(\cdot, \tilde{z}) - \delta/2$. This earns $\delta/2$ on each type served. Let $\hat{\theta}_l$ be the lowest type with $\hat{\theta}_l \geq \theta_l$ who accepts versus $\bar{v}(\hat{\theta}_l, z_l)$, and let $\hat{\theta}_h$ be the highest type with $\hat{\theta}_h \leq \theta_h$ who accepts versus $\bar{v}(\hat{\theta}_h, z_h)$. Since $\bar{v}(\cdot, z_l)$ is an upper bound on the surplus that z_l (if such a firm even exists) will offer in any post-entry equilibrium, and similarly for $\bar{v}(\cdot, z_h)$, any type between $\hat{\theta}_l$ and $\hat{\theta}_h$ is certainly attracted to the entrant. Note in particular that since V is strictly supermodular, $\bar{v}(\cdot, \tilde{z})$ is steeper than $\bar{v}(\cdot, z_l)$, and so no type above $\tilde{\theta}$ prefers any firm at or below z_l to $\tilde{z}$, and similarly, no type below $\tilde{\theta}$ prefers any firm at or above z_h to $\tilde{z}$.

If $\hat{\theta}_l = \theta_l$ or $\hat{\theta}_h = \theta_h$ then $\hat{\theta}_h - \hat{\theta}_l \geq (\theta_h - \theta_l)/2$, and the firm earns at least $((\theta_h - \theta_l)/2)\ell_h(\delta/2) \geq (\ell_h \ell_{\bar{a}_z} \ell_{z^{\theta}_{\theta}} / 32 v^3_{z^{\theta}_{\theta}}) d_z^3$, using (20) and the definition of δ. Otherwise, $\hat{\theta}_l$ is defined by $\bar{v}(\hat{\theta}_l, \tilde{z}) - \bar{v}(\hat{\theta}_l, z_l) = \delta/2$. But, $\bar{v}(\tilde{\theta}, \tilde{z}) - \bar{v}(\tilde{\theta}, z_l) = v_*(\tilde{\theta}) - \bar{v}(\tilde{\theta}, z_l) \geq \delta$ by (21), and so $(\bar{v}(\tilde{\theta}, \tilde{z}) - \bar{v}(\tilde{\theta}, z_l)) - (\bar{v}(\hat{\theta}_l, \tilde{z}) - \bar{v}(\hat{\theta}_l, z_l)) \geq \delta/2$, from which, since $(\bar{v}(\theta, \tilde{z}) - \bar{v}(\theta, z_l))_{\theta} = \bar{a}(\theta, \tilde{z}) - \bar{a}(\theta, z_l) \leq (\tilde{z} - z_l)\, v_{\bar{a}_z}$, we have $\tilde{\theta} - \hat{\theta}_l \geq \delta / (2(\tilde{z} - z_l)\, v_{\bar{a}_z})$, and similarly for $\hat{\theta}_h - \tilde{\theta}$, and thus

$$\hat{\theta}_h - \hat{\theta}_l \geq \frac{\delta}{2 v_{\bar{a}_z}} \left(\frac{1}{\tilde{z} - z_l} + \frac{1}{z_h - \tilde{z}} \right) = \frac{\delta}{2 v_{\bar{a}_z}} \frac{z_h - z_l}{(\tilde{z} - z_l)(z_h - \tilde{z})} \geq \frac{\delta}{2 v_{\bar{a}_z}} \frac{4}{d_z},$$

where the second inequality follows since $(\tilde{z} - z_l)(z_h - \tilde{z}) \leq (z_h - z_l)^2/4$, and so, using (21) the firm earns at least $(\hat{\theta}_h - \hat{\theta}_l)\ell_h \delta/2 \geq \delta^2 \ell_h / (v_{\bar{a}_z} d_z) = (\ell^2_{\bar{a}_z} \ell^2_{z^{\theta}_{\theta}} \ell_h / 64 v^4_{z^{\theta}_{\theta}} v_{\bar{a}_z}) d_z^3$. Thus, we have established that an entrant can earn at least $\rho_1 d_z^3$, where $\rho_1 = \min\left\{ \ell_{\bar{a}_z} \ell_{z^{\theta}_{\theta}} \ell_h / 32 v^3_{z^{\theta}_{\theta}}, (\ell^2_{\bar{a}_z} \ell^2_{z^{\theta}_{\theta}} \ell_h / 64 v^4_{z^{\theta}_{\theta}} v_{\bar{a}_z}) \right\}$.

Step 2 Let us next show that for each firm $1 < n < N$, there is an upper bound on how much Firm n can earn as function of how far apart its competitors are. In particular, we claim there is $\rho_2 \in (0, \infty)$ such that each such firm earns at most $\rho_2 \left(z^{n+1} - z^{n-1}\right)^3$.

Fix $1 < n < N$, let n serve $[\theta_l, \theta_h]$, choose $\hat{\theta} \in [\theta_l, \theta_h]$ and let $\hat{a} = \alpha^n(\hat{\theta})$. Our first task is to show that

$$V^n(\hat{a}) - \max_{n'} V^{n'}(\hat{a}) \leq \frac{-\ell_{V_{zz}}}{2}\left(z^{n+1} - z^{n-1}\right)^2. \tag{22}$$

If $z^n = z^{n+1}$, then this holds trivially, since the lhs is zero. So, assume $z^{n+1} > z^n$. Recall by NP that $V(\hat{a}, z^n) \geq V\left(\hat{a}, z^{n+1}\right)$. Let $\hat{z} = z^a(\hat{a}) = \arg\max_z V(\hat{a}, z)$ be the technology that is most efficient at $\hat{a}$. We claim that $\hat{z} \in \left[z^{n-1}, z^{n+1}\right]$. To see this, assume $\hat{z} > z^{n+1}$ (the case $\hat{z} < z^{n-1}$ is similar). Then, since $z^{n+1} > z^n$, there is $\alpha \in (0,1)$ such that $z^{n+1} = \alpha z^n + (1-\alpha)\hat{z}$. But, then, since V is strictly concave in z, $V\left(\hat{a}, z^{n+1}\right) > \alpha V(\hat{a}, z^n) + (1-\alpha)V(\hat{a}, \hat{z}) > V\left(\hat{a}, z^{n+1}\right)$, a contradiction. But then, since $V_z(\hat{a}, \hat{z}) = 0$, we have by Taylor's Theorem that

$$V^n(\hat{a}) - \max_{n'} V^{n'}(\hat{a}) \leq V(\hat{a}, \hat{z}) - V(\hat{a}, z^{n-1}) \leq \frac{-\ell_{V_{zz}}}{2}\left(\hat{z} - z^{n-1}\right)^2 \leq \frac{-\ell_{V_{zz}}}{2}\left(z^{n+1} - z^{n-1}\right)^2,$$

where the last inequality follows since $\hat{z} \in \left[z^{n-1}, z^{n+1}\right]$, and we have established (22).

Recall from Section 4.1.2 that $\pi^n(\hat{\theta}, \hat{a}, v^n) \leq V^n(\hat{a}) - \max_{n' \neq n} V^{n'}(\hat{a})$. Thus, using (22), the total profit of Firm n is at most $-\ell_{V_{zz}} \upsilon_h \left(z^{n+1} - z^{n-1}\right)^2 (\theta_h - \theta_l)/2$,and it is enough to show that $(\theta_h - \theta_l)$ is bounded by a multiple of $\left(z^{n+1} - z^{n-1}\right)$. Let $a^z = (z^a)^{-1}$, so that at any z, $a^z(z)$ is the action that z is uniquely best at providing. But, since $[\alpha^n(\theta_l), \alpha^n(\theta_h)] \subseteq [a_e^{n-1}, a_e^n] \subseteq (a^z\left(z^{n-1}\right), a^z\left(z^{n+1}\right))$, and since $\alpha^n = \gamma^n(\cdot, \kappa^n)$, we have $\gamma^n(\theta_h, \kappa^n) - \gamma^n(\theta_l, \kappa^n) \leq a^z\left(z^{n+1}\right) - a^z\left(z^{n-1}\right)$, and so $\theta_h - \theta_l \leq (\upsilon_{a_z^z}/\ell_{\gamma_\theta^n})\left(z^{n+1} - z^{n-1}\right)$, where $\ell_{\gamma_\theta^n} > 0$ is taken over $\theta \in [0,1]$, and $\kappa^n \in [0,1]$, and $\upsilon_{a_z^z} < \infty$ is taken over the compact subset $z^a([0, \tilde{a}])$. Thus, the total profit of Firm n is at most $\rho_2\left(z^{n+1} - z^{n-1}\right)^3$, where $\rho_2 = \left(-\ell_{V_{zz}} \upsilon_h \upsilon_{a_z^z}\right)/2\ell_{\gamma_\theta^n}$, and we are done.

Step 3 There is ρ such that $1/(\rho F^{1/3}) \leq N \leq (\rho/F^{1/3}) + 2$, and d_z is $O(1/N)$.

This follows from the last two steps since no firms wish to enter or exit. For entry not to be profitable, we must have $\rho_1 d_z^3 \leq F$, and thus $d_z \leq (F/\rho_1)^{1/3}$, and hence, since $N \geq \left(z^\theta(1) - z^\theta(0)\right)/d_z$, we have $N \geq \left(z^\theta(1) - z^\theta(0)\right)(\rho_1)^{1/3}/F^{1/3}$. Similarly, for firm $1 < n < N$ not to want to exit, we must have $F \leq \rho_2\left(z^{n+1} - z^{n-1}\right)^3$, and so $(F/\rho_2)^{1/3} \leq \left(z^{n+1} - z^{n-1}\right)$. But, then

$$(N-2)\left(\frac{F}{\rho_2}\right)^{\frac{1}{3}} \leq \sum_{n=2}^{N-1}\left(z^{n+1} - z^{n-1}\right) = \sum_{n=2}^{N-1}\left(z^{n+1} - z^n\right) + \sum_{n=2}^{N-1}\left(z^n - z^{n-1}\right) \leq 2\bar{z},$$

and so, rearranging the end terms, $N \leq (2\bar{z}\rho_2^{1/3}/F^{1/3}) + 2$,and so, taking ρ large enough, we have established the first claim. That d_z is $O(1/N)$ follows immediately, since $d_z \leq (F/\rho_1)^{1/3}$.

Step 4 Let $d_1 = \max_{a \in \left[a^\theta(0), a^\theta(1)\right], n}(V^n(a) - \max_{n' \neq n} V^{n'}(a)) \geq 0$ be the largest difference over the efficient range of actions between the first and second highest V. We claim that there is ρ_3 such that $d_1 \leq \rho_3 d_z^2$ and hence, by Step 3, d_1 is of order $1/N^2$.

This is much as in Step 2. Fix $\hat{a} \in [0, \tilde{a}]$, and let $\hat{z} = z^a(\hat{a})$. Let z be any other technology.

Then, much as in (22),

$$V(\hat{a},\hat{z})-V(\hat{a},z)\leq-\frac{1}{2}(\hat{z}-z)^2\,\ell_{V_{zz}}. \tag{23}$$

Now, assume that $\hat{a}\in\left[a^{\theta}(0),a^{\theta}(1)\right]$, so that $\hat{z}\in\left[z^{\theta}(0),z^{\theta}(1)\right]$. Pick the two operating firms closest to $\hat{z}$. Each is at most $2d_z$ away from $\hat{z}$. Then, the first most efficient such firm has V at most $V(\hat{a},\hat{z})$, and the second has V at least $V(\hat{a},\hat{z})+\frac{1}{2}(2d_z)^2\,\ell_{V_{zz}}$, and so $d_1\leq-2\ell_{V_{zz}}d_z^2$, and so, we can take $\rho_3=-2\ell_{V_{zz}}$.

Step 5 Each type θ must in equilibrium receive an amount close to $v_*(\theta)=\max_{a,z}(V(a,z)+\theta a)$. In particular, we claim that when $d_z<\left(z^{\theta}(1)-z^{\theta}(0)\right)/4$, then there is ρ_4 such that θ's equilibrium payoff is at least $v_*(\theta)-\rho_4 d_z^2$, and hence by Step 3, the difference between θ's equilibrium payoff and $v_*(\theta)$ is of order $1/N^2$.

The idea here is that since firms are densely packed on $\left[z^{\theta}(0),z^{\theta}(1)\right]$, there are at least two firms that are very well-positioned to meet the needs of any given type. Competition and incentive compatibility then force the equilibrium payoff to be near v_*. Formally, consider any type $\tilde{\theta}\in[0,1]$. By definition of d_z, there is a firm n' for whom $z^{n'}$ is within d_z of $z^{\theta}(\tilde{\theta})$, and hence a firm n with $z^{\theta}(0)\leq z^{n-1}\leq z^{n+1}\leq z^{\theta}(1)$ for whom z^n is within $2d_z$ of $z^{\theta}(\tilde{\theta})$. Let $\hat{\theta}$ be any type n serves, and let $\hat{a}$ be the associated action. Let us show that $\tilde{\theta}$ can attain at least $v_*(\theta)-\rho_4 d_z^2$ for a suitable ρ_4 by imitating $\hat{\theta}$.

Note first that since $\hat{a}\in[a_e^{n-1},a_e^n]\subseteq\left[a^z\left(z^{n-1}\right),a^z\left(z^{n+1}\right)\right]$,

$$\hat{a}-a^{\theta}(\tilde{\theta})=\hat{a}-a^z(z^{\theta}(\tilde{\theta}))\leq a^z\left(z^{n+1}\right)-a^z(z^{\theta}(\tilde{\theta}))\leq\left|z^{n+1}-z^{\theta}(\tilde{\theta})\right|\upsilon_{a_z^z}\leq 3d_z\upsilon_{a_z^z}.$$

Take a first-order Taylor expansion (Mardsen and Tromba (2012), Theorem 2, pp.160–162) of $V(a,z)+a\tilde{\theta}$ as a function of a and z around $(a^{\theta}(\tilde{\theta}),z^{\theta}(\tilde{\theta}))$, noting that the first order first-order terms disappear since $(a^{\theta}(\tilde{\theta}),z^{\theta}(\tilde{\theta}))$ is a maximum of $V(a,z)+a\tilde{\theta}$. Then, using that $v_*(\tilde{\theta})=V(a^{\theta}(\tilde{\theta}),z^{\theta}(\tilde{\theta}))+a^{\theta}(\tilde{\theta})\tilde{\theta}$, write

$$V(\hat{a},z^n)+\hat{a}\tilde{\theta}=v_*(\tilde{\theta})+\frac{V_{aa}(a',z')}{2}(\hat{a}-a^{\theta}(\tilde{\theta}))^2+\frac{V_{zz}(a',z')}{2}(z^n-z^{\theta}(\tilde{\theta}))^2+V_{az}(a',z')(\hat{a}-a^{\theta}(\tilde{\theta}))(z^n-z^{\theta}(\tilde{\theta})),$$

where (a',z') is some point on the line segment between $(\hat{a},z^n)$ and $(a^{\theta}(\tilde{\theta}),z^{\theta}(\tilde{\theta}))$. Hence

$$V(\hat{a},z^n)+\hat{a}\tilde{\theta}\geq v_*(\tilde{\theta})+\frac{\ell_{V_{aa}}}{2}(3d_z\upsilon_{a_z^z}))^2+\frac{\ell_{V_{zz}}}{2}(2d_z)^2-\upsilon_{V_{az}}6\upsilon_{a_z^z}d_z^2,$$

which, since $d_1\leq\rho_3 d_z^2$ implies that the value to $\tilde{\theta}$ of imitating $\hat{\theta}$ is at least $V(\hat{a},z^n)+\hat{a}\tilde{\theta}-d_1\geq v_*(\tilde{\theta})-\rho_4 d_z^2$, where $\rho_4=(9/2)\ell_{V_{aa}}\upsilon_{a_z^z}^2+2\ell_{V_{zz}}-\upsilon_{V_{az}}6\upsilon_{a_z^z}-\rho_3$.

Step 6 The firm serving $\theta=1$ is not very far above $z^{\theta}(1)$, and the firm serving $\theta=0$ is not far below $z^{\theta}(0)$. Indeed, there is ρ_5 such that each difference is at most $\rho_5 d_z$.

The idea is that when z is much above $z^{\theta}(1)$, then a firm with technology z, even if it offers all available surplus to the agent, is unable to beat the offerings of firms near $z^{\theta}(1)$. To begin, for given $\hat{z} > z^{\theta}(1)$, let us bound $\bar{v}(1,\hat{z})$, the most surplus $\hat{z}$ can offer type 1. A similar argument will apply for $\hat{z} < z^{\theta}(0)$. But, $\bar{v}\left(1, z^{\theta}(1)\right) = v^{*}(1)$, while $\bar{v}_{z}\left(1, z^{\theta}(1)\right) = 0$, and so, taking a first-order Taylor Expansion of $\bar{v}(1,\cdot)$ around $z^{\theta}(1)$, for some $\tilde{z} \in [z^{\theta}(1), \hat{z}]$,

$$\bar{v}(1,\hat{z}) = v^{*}(1) + \left(\hat{z} - z^{\theta}(1)\right)^{2} \bar{v}_{zz}(1,\tilde{z})/2 \leq v^{*}(1) + \left(\hat{z} - z^{\theta}(1)\right)^{2} \upsilon_{\bar{v}_{zz}(1,\cdot)}/2.$$

Let us show next that $\upsilon_{\bar{v}_{zz}(1,\cdot)} < 0$. By the Envelope Theorem, $\bar{v}_{z}(1,z) = \left(V(\bar{a}(1,z),z) + \bar{a}(1,z)\right)_{z} = V_{z}(\bar{a}(1,z),z)$, and so

$$\begin{aligned}\bar{v}_{zz}(1,z) &= V_{za}(\bar{a}(1,z),z)\,\bar{a}_{z}(1,z) + V_{zz}(\bar{a}(1,z),z) = -V_{za}(\bar{a}(1,z),z)\frac{V_{za}(\bar{a}(1,z),z)}{V_{aa}(\bar{a}(1,z),z)} + V_{zz}(\bar{a}(1,z),z)\\ &\leq \max_{[0,\tilde{a}]\times[0,z]}\left(\frac{1}{V_{aa}(a,z)}\left(V_{aa}(a,z)\,V_{zz}(a,z) - V_{za}^{2}(a,z)\right)\right) = \upsilon_{\bar{v}_{zz}(1,\cdot)} < 0,\end{aligned}$$

where the second inequality follows by the concavity assumptions on V and because $[0,\tilde{a}]\times[0,z]$ is compact.

To complete this step, assume that $\hat{z} > z^{\theta}(1)$ serves $\theta = 1$. Then, since type 1 earns at least $v_{*}(1) - \rho_{4}d_{z}^{2}$ by dealing with a firm with $z \cong z^{\theta}(1)$, we have by PP that $v_{*}(1) + \upsilon_{\bar{v}_{zz}(1,\cdot)}\left(\hat{z} - z^{\theta}(1)\right)^{2}/2 \geq v_{*}(1) - \rho_{4}d_{z}^{2}$, and so, $\hat{z} - z^{\theta}(1) \leq \sqrt{(-2\rho_{4})/\upsilon_{\bar{v}_{zz}(1,\cdot)}}\,d_{z} \equiv \rho_{5}d_{z}$.

Step 7 The profits on each type are bounded by $\rho_{6}d_{z}^{2}$, and so of order $1/N^{2}$.

This follows since, by the preceding steps, for each equilibrium firm, type, and action, there is a nearby firm of similar capabilities, since by Step 6, there is no firm very far outside of $\left[z^{\theta}(0), z^{\theta}(1)\right]$, and by the preceding steps, firms are closely packed in $\left[z^{\theta}(0), z^{\theta}(1)\right]$. Formally, fix $\hat{\theta}$, and assume that $\hat{\theta}$ receives action $\hat{a}$ in equilibrium. If $\hat{a} \in \left[a^{\theta}(0), a^{\theta}(1)\right]$, then by Step 4, we have $\pi^{n}(\hat{\theta},\hat{a},v^{n}) \leq d_{1} \leq \rho_{3}d_{z}^{2}$, where the first inequality is as proven in Step 2. If $\hat{a} \notin \left[a^{\theta}(0), a^{\theta}(1)\right]$ then by Step 6, the firm serving $\hat{\theta}$ has $\hat{z}$ within $\rho_{5}d_{z}$ of $\left[z^{\theta}(0), z^{\theta}(1)\right]$, and so $\hat{z}$ is within $(\rho_{5}+1)\,d_{z}$ of $\tilde{z}$, where $\tilde{z} \in \left[z^{\theta}(0), z^{\theta}(1)\right]$ is some other operating firm. But then, again as in Step 2, and using (23),

$$\pi^{n}(\hat{\theta},\hat{a},v^{n}) \leq V(\hat{a},\hat{z}) - V(\hat{a},\tilde{z}) \leq -\frac{1}{2}(\hat{z}-\tilde{z})^{2}\,\ell_{V_{zz}} \leq -\frac{1}{2}\ell_{V_{zz}}(\rho_{5}+1)^{2}d_{z}^{2},$$

and so defining $\rho_{6} = \max\{\rho_{3}, -(1/2)\,\ell_{V_{zz}}(\rho_{5}+1)^{2}\}$, we are done, noting that from Step 3, d_{z} is $O(1/N)$, and hence π^{n} is $O\left(1/N^{2}\right)$.

5 Omitted Proofs for Section 11

Lemma 24 *Consider* γ *and* $\tilde{\kappa}$ *as functions on* $\tilde{R} \cap \Theta$*. Then,* $(\gamma(\theta_l, \tilde{\kappa}))_{\theta_l} > \gamma_\kappa(\theta_l, \tilde{\kappa})\tilde{\kappa}_{\theta_l} > 0$ *and* $(\gamma(\theta_h, \tilde{\kappa}))_{\theta_h} > \gamma_\kappa(\theta_h, \tilde{\kappa})\tilde{\kappa}_{\theta_h} > 0$*, with*

$$\begin{vmatrix} (\gamma(\theta_l, \tilde{\kappa}))_{\theta_l} & (\gamma(\theta_l, \tilde{\kappa}))_{\theta_h} \\ (\gamma(\theta_h, \tilde{\kappa}))_{\theta_l} & (\gamma(\theta_h, \tilde{\kappa}))_{\theta_h} \end{vmatrix} > 0.$$

Proof Note that $(\gamma(\theta_l, \tilde{\kappa}))_{\theta_l} = \gamma_\theta(\theta_l, \tilde{\kappa}) + \gamma_\kappa(\theta_l, \tilde{\kappa})\tilde{\kappa}_{\theta_l} > \gamma_\kappa(\theta_l, \tilde{\kappa})\tilde{\kappa}_{\theta_l}$, since $\gamma_\theta > 0$ using that $\tilde{\kappa} \in [0, 1]$. But, since $\iota(\theta_l, \theta_h, \tilde{\kappa}) = 0$ on Θ, we have $\tilde{\kappa}_{\theta_l} = -\iota_{\theta_l}/\iota_\kappa < 0$, since $\iota_\theta > 0$ and $\iota_\kappa > 0$ using the definition of ι. Thus $\gamma_\kappa(\theta_l, \tilde{\kappa})\tilde{\kappa}_{\theta_l} > 0$, since $\gamma_\kappa < 0$. Similarly, $\tilde{\kappa}_{\theta_h} < 0$, and so $(\gamma(\theta_h, \tilde{\kappa}))_{\theta_h} > \gamma_\kappa(\theta_h, \tilde{\kappa})\tilde{\kappa}_{\theta_h} > 0$. But then,

$$\begin{vmatrix} (\gamma(\theta_l, \tilde{\kappa}))_{\theta_l} & (\gamma(\theta_l, \tilde{\kappa}))_{\theta_h} \\ (\gamma(\theta_h, \tilde{\kappa}))_{\theta_l} & (\gamma(\theta_h, \tilde{\kappa}))_{\theta_h} \end{vmatrix} > \begin{vmatrix} \gamma_\kappa(\theta_l, \tilde{\kappa})\tilde{\kappa}_{\theta_l} & \gamma_\kappa(\theta_l, \tilde{\kappa})\tilde{\kappa}_{\theta_h} \\ \gamma_\kappa(\theta_h, \tilde{\kappa})\tilde{\kappa}_{\theta_l} & \gamma_\kappa(\theta_h, \tilde{\kappa})\tilde{\kappa}_{\theta_h} \end{vmatrix} = 0.\square$$

Proof of Lemma 14 From (12), and recalling that π_a does not depend on $\tilde{v}$,

$$\begin{aligned} \frac{r_{\theta_h\theta_l}(\theta_l, \theta_h)}{h(\theta_h)} &= \pi_{aa}(\theta_h, \gamma(\theta_h, \tilde{\kappa}), \tilde{v}) \, (\gamma(\theta_h, \tilde{\kappa}))_{\theta_l} \left(a^{-n}(\theta_h) - \gamma(\theta_h, \tilde{\kappa})\right) + \pi_a(\theta_h, \gamma(\theta_h, \tilde{\kappa}), \tilde{v})(-\left(\gamma(\theta_h, \tilde{\kappa})\right)_{\theta_l}) \\ &\quad + \pi_a(\theta_h, \gamma(\theta_h, \tilde{\kappa}), \tilde{v}) \, (\gamma(\theta_h, \tilde{\kappa}))_{\theta_l}, \\ &= \pi_{aa}(\theta_h, \gamma(\theta_h, \tilde{\kappa}), \tilde{v}) \, (\gamma(\theta_h, \tilde{\kappa}))_{\theta_l} \left(a^{-n}(\theta_h) - \gamma(\theta_h, \tilde{\kappa})\right), \end{aligned} \tag{24}$$

and similarly, from (13),

$$\frac{r_{\theta_l\theta_h}(\theta_l, \theta_h)}{h(\theta_l)} = \pi_{aa}(\theta_l, \gamma(\theta_l, \tilde{\kappa}), \tilde{v}) \, (\gamma^n(\theta_l, \tilde{\kappa}))_{\theta_h} \left(\gamma(\theta_l, \tilde{\kappa}) - a^{-n}(\theta_l)\right). \tag{25}$$ [49]

To see that $r_{\theta_l\theta_h} < 0$, start from (25) (or analogously from (24)), and note that $\pi_{aa} < 0$, that $(\gamma^n(\theta_l, \tilde{\kappa}))_{\theta_h} = \gamma_\kappa(\theta_l, \tilde{\kappa})\tilde{\kappa}_{\theta_h} > 0$, and that by stacking, $\gamma(\theta_l, \tilde{\kappa}) - a^{-n}(\theta_l) > 0$.

Note next that since $\iota(\theta_l, \theta_h, \tilde{\kappa}) = 0$, $\tilde{v}(\theta_l) = v^{-n}(\theta_l)$ and $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$. Note that $\pi_a(\theta_h, \gamma(\cdot, \tilde{\kappa}), \tilde{v}) =_s \tilde{\kappa} - H(\theta_h) \le 0$, since $\tilde{\kappa} \in [H(\theta_l), H(\theta_h)]$. Similarly, $\pi_a(\theta_l, \gamma(\cdot, \tilde{\kappa}), \tilde{v}) \ge 0$. Assume that $r_{\theta_h}(\theta_l, \theta_h) = 0$. Then using (12),

$$\begin{aligned} \frac{r_{\theta_h\theta_h}(\theta_l, \theta_h)}{h(\theta_h)} = &\left(1 + \pi_{aa}(\theta_h, \gamma(\theta_h, \tilde{\kappa}), \tilde{v}) \, (\gamma(\theta_h, \tilde{\kappa}))_{\theta_h}\right) \left(a^{-n}(\theta_h) - \gamma(\theta_h, \tilde{\kappa})\right) \\ &+ \pi_a(\theta_h, \gamma(\theta_h, \tilde{\kappa}), \tilde{v})(a_\theta^{-n}(\theta_h) - (\gamma(\theta_h, \tilde{\kappa}))_{\theta_h}) \\ &+ \gamma(\theta_h, \tilde{\kappa}) + \pi_a(\theta_h, \gamma(\theta_h, \tilde{\kappa}), \tilde{v}) \, (\gamma(\theta_h, \tilde{\kappa}))_{\theta_h} - a^{-n}(\theta_h) \end{aligned}$$

[49] These two expressions must of course be equal, but it is convenient to express them in these two different ways.

where the term involving h_θ disappears since $r_{\theta_h} = 0$, and where we use that $\tilde{v}(\theta_h) = v^{-n}(\theta_h)$, and hence $(\tilde{v}(\theta_h))_{\theta_h} = (v^{-n}(\theta_h))_{\theta_h} = a^{-n}(\theta_h)$. Cancelling,

$$\begin{aligned}
\frac{r_{\theta_h\theta_h}(\theta_l,\theta_h)}{h(\theta_h)} &= \pi_{aa}(\theta_h,\gamma(\theta_h,\tilde{\kappa}),\tilde{v})\,(\gamma(\theta_h,\tilde{\kappa}))_{\theta_h}\left(a^{-n}(\theta_h)-\gamma(\theta_h,\tilde{\kappa})\right)+\pi_a(\theta_h,\gamma(\theta_h,\tilde{\kappa}),\tilde{v})a_\theta^{-n}(\theta_h) \\
&\le \pi_{aa}(\theta_h,\gamma(\theta_h,\tilde{\kappa}),\tilde{v})\,(\gamma(\theta_h,\tilde{\kappa}))_{\theta_h}\left(a^{-n}(\theta_h)-\gamma(\theta_h,\tilde{\kappa})\right) \qquad (26)\\
&< 0,
\end{aligned}$$

where the first inequality uses that $\pi_a(\theta_h,\gamma(\theta_h,\tilde{\kappa}),\tilde{v}) \le 0$ and the second uses that $\pi_{aa} < 0$, that by Lemma 24, $(\gamma(\theta_h,\tilde{\kappa}))_{\theta_h} > 0$, and that by stacking, $C1$, and $\tilde{\kappa} \in [0,1]$, $a^{-n}(\theta_h) - \gamma(\theta_h,\tilde{\kappa}) > 0$.[50]

Similarly, taking cancellations as before, if $r_{\theta_l} = 0$, then

$$\frac{r_{\theta_l\theta_l}(\theta_l,\theta_h)}{h(\theta_l)} \le \pi_{aa}(\theta_l,\gamma(\theta_l,\tilde{\kappa}),\tilde{v})\,(\gamma(\theta_l,\tilde{\kappa}))_{\theta_l}\left(\gamma(\theta_l,\tilde{\kappa})-a^{-n}(\theta_l)\right) < 0. \qquad (27)$$

For strict local concavity, it remains to show that if $r_{\theta_l} = r_{\theta_h} = 0$, then $r_{\theta_l\theta_l}r_{\theta_h\theta_h} - r^2_{\theta_l\theta_h} > 0$. From (26) and (27),

$$\begin{aligned}
\frac{r_{\theta_l\theta_l}r_{\theta_h\theta_h}}{h(\theta_l)h(\theta_h)} \ge\ & \pi_{aa}(\theta_h,\gamma(\theta_h,\tilde{\kappa}),\tilde{v})\,(\gamma(\theta_h,\tilde{\kappa}))_{\theta_h}\left(a^{-n}(\theta_h)-\gamma(\theta_h,\tilde{\kappa})\right)\\
&\times \pi_{aa}(\theta_l,\gamma(\theta_l,\tilde{\kappa}),\tilde{v})\,(\gamma(\theta_l,\tilde{\kappa}))_{\theta_l}\left(\gamma(\theta_l,\tilde{\kappa})-a^{-n}(\theta_l)\right),
\end{aligned}$$

while from (24) and (25),

$$\begin{aligned}
\frac{r_{\theta_l\theta_h}r_{\theta_h\theta_l}}{h(\theta_l)h(\theta_h)} =\ & \pi_{aa}(\theta_l,\gamma(\theta_l,\tilde{\kappa}),\tilde{v})\,(\gamma(\theta_l,\tilde{\kappa}))_{\theta_h}\left(\gamma(\theta_l,\tilde{\kappa})-a^{-n}(\theta_l)\right)\\
&\times \pi_{aa}(\theta_h,\gamma(\theta_h,\tilde{\kappa}),\tilde{v})\,(\gamma(\theta_h,\tilde{\kappa}))_{\theta_l}\left(a^{-n}(\theta_h)-\gamma(\theta_h,\tilde{\kappa})\right).
\end{aligned}$$

Collecting common positive terms, it suffices that $(\gamma(\theta_h,\tilde{\kappa}))_{\theta_h}(\gamma(\theta_l,\tilde{\kappa}))_{\theta_l} - (\gamma(\theta_l,\tilde{\kappa}))_{\theta_h}(\gamma(\theta_h,\tilde{\kappa}))_{\theta_l} > 0$, which follows from Lemma 24. □

[50] To be careful, a_θ^{-n}, and hence $r_{\theta_h\theta_h}$, may not be everywhere defined. But, since a^{-n} is increasing, $\liminf_{\varepsilon\downarrow 0} a_\theta^{-n}(\theta_h+\varepsilon) \ge 0$ and $\liminf_{\varepsilon\downarrow 0} a_\theta^{-n}(\theta_h-\varepsilon) \ge 0$, and so, since $\pi_a(\theta_h,\gamma(\theta_h,\tilde{\kappa}),\tilde{v}) \le 0$, we have $\limsup_{\varepsilon\downarrow 0} r_{\theta_h\theta_h}(\theta_l,\theta_h+\varepsilon) < 0$, and $\limsup_{\varepsilon\downarrow 0} r_{\theta_h\theta_h}(\theta_l,\theta_h-\varepsilon) < 0$. We henceforth ignore this technical detail.