# Equilibrium Uniqueness in Entry Games with Combinatorial Actions and Endogenous Pricing*

Hyungjin Kim
Korea Development Institute

April 4, 2026

**Abstract**

I prove equilibrium uniqueness for a class of combinatorial entry games with endogenous pricing—games in which each firm chooses a *subset* of available actions (e.g., which countries to produce in) and firms then compete under variable markups à la Atkeson and Burstein (2008). The combination of combinatorial entry and endogenous markups creates a three-way circularity—activation shifts market shares, which change markups, which feed back into activation incentives—that existing frameworks do not address. I show this circularity is self-limiting: when firms have private information on action-specific fixed costs, the activation decision reduces to a threshold system whose contraction modulus decomposes into three interpretable forces. The modulus vanishes at both extremes of noise—large noise dampens responses while small noise pushes thresholds into the tails of the cost distribution—so when its peak is below one the Bayesian Nash equilibrium is unique for *all* noise levels—a condition verifiable from model primitives without solving the game. Each firm's activation problem remains supermodular despite the pricing channel, so existing combinatorial solvers apply without modification. In numerical experiments on multinational production-site activation with up to $2^{40}$ candidate subsets per firm, the peak contraction modulus is 0.95, confirming global uniqueness.

*Keywords:* entry games, equilibrium uniqueness, combinatorial discrete choice, Bayesian Nash equilibrium, variable markups, supermodular games
*JEL Codes:* C72, C62, C63, L13, F12

*I thank Jaedo Choi, Sungwan Hong, Ryo Makioka, and Hongyong Zhang for helpful discussions.
E-mail: hyungjin.kim@kdi.re.kr.

# 1. Introduction

Entry games are a workhorse of industrial organization and international trade, yet they are plagued by equilibrium multiplicity: when entry decisions are strategic substitutes—one firm's entry reduces rivals' incentives to enter—multiple configurations of active firms can be mutually consistent. Espin-Sanchez et al. (2023) show that when firms have private information over entry costs, the resulting Bayesian game has a unique equilibrium for games with *binary* entry decisions and scalar private types. Many economically important settings, however, involve *combinatorial* entry: each player chooses which *subset* of available actions to activate—for example, a multinational firm choosing which countries to produce in—and payoffs exhibit within-player complementarities across actions. Endogenous pricing adds a further complication: activating an action changes market structure and hence all players' markups. The resulting three-way circularity between activation, market shares, and pricing creates new obstacles to uniqueness that binary-entry results do not address.

This paper proves equilibrium uniqueness for a class of combinatorial entry games with endogenous pricing. Each player selects a subset of available actions—formalized as a *combinatorial discrete choice problem* (CDCP) in the sense of Jia (2008) and Arkolakis et al. (2023)—and players then compete in quantities under CES demand with endogenous variable markups à la Atkeson and Burstein (2008, hereafter AB). Neither the entry-game literature nor the CDCP literature has incorporated endogenous pricing: Espin-Sanchez et al. (2023) assume exogenous payoff functions, while the CDCP literature assumes constant markups. I bridge this gap and show that endogenous pricing, despite introducing a new source of strategic interaction, does not generate multiplicity at empirically relevant parameterizations—and where multiplicity does arise at extreme parameters, the Bayesian extension with private costs selects a unique equilibrium.

The leading application is multinational firms' production-site activation, where each firm chooses which subset of countries to produce in and firms compete in quantities in segmented international markets. This setting motivates the CES demand structure and the specific form of the pricing game, and provides foundations for dynamic models of multinational production. The game-theoretic results, however, apply broadly to any combinatorial entry game with contractive pricing—for example, retail chain location (Jia,

2008), platform market entry, or technology adoption with network effects.

I make four contributions. First, I establish sharp conditions for equilibrium uniqueness on both the intensive and extensive margins. On the intensive margin, I derive a market-share threshold for pricing contraction based on the interaction matrix of the markup game (Proposition 11 and Corollary 17). On the extensive margin, when firms have private information on action-specific fixed costs—an empirically natural feature of entry games—the combinatorial activation decision becomes a continuous threshold problem whose Jacobian decomposes into three interpretable forces: density dampening, marginal-value sensitivity, and threshold location (Proposition 13). The resulting contraction modulus has a "bathtub" shape—vanishing at both extremes of the noise scale (at large noise, density dampening dominates; at small noise, all thresholds are pushed to the tails of the distribution)—so that when the complete-information equilibrium is unique and the peak contraction modulus is below one, the Bayesian Nash equilibrium (BNE) is unique for *all* noise levels.

Second, I show that each player's activation problem is *supermodular* despite the countervailing pricing channel, extending Antràs et al.'s (2017) constant-markup result to endogenous variable markups (Proposition 6). Supermodularity enables lattice-theoretic algorithms to solve each player's combinatorial problem without exhaustive search, and gives the threshold system the monotone structure required for the uniqueness result.

Third, I characterize the equilibrium set under strategic substitutes across players. The extremal fixed points of the squared best-response map bound the entire Nash equilibrium set (Proposition 8), and iterated best response from the supremum converges monotonically to these bounds (Proposition 9).

Fourth, I provide a systematic comparison of four algorithm variants—Jia–Tarski vs. AES inner solvers crossed with Jacobi vs. Gauss–Seidel outer iteration—each nesting the endogenous-markup fixed point as an innermost layer. Numerical experiments with up to $2^{40}$ candidate subsets per player confirm uniqueness across all tested parameterizations. A comparison against constant CES markups shows that endogenous pricing expands activation by 16% and raises aggregate profits by 30%, underscoring the quantitative importance of the pricing channel.

**Entry games and equilibrium selection.** This paper contributes primarily to the literature on equilibrium uniqueness in entry games. Beginning with Bresnahan and Reiss (1991) and Berry (1992), the empirical entry literature has confronted multiplicity as a barrier to estimation: Berry restricts equilibrium selection to obtain a tractable likelihood; Ciliberto and Tamer (2009) use partial identification to bound parameters without selecting among equilibria; Seim (2006) introduces endogenous product-type choices, a structure related to combinatorial entry; Igami and Yang (2016) estimate a dynamic entry model with combinatorial product portfolios using simulation methods that must navigate the multiplicity problem. Espin-Sanchez et al. (2023) prove BNE uniqueness in static entry games where each firm makes a *binary* enter/not-enter decision with scalar private types; their threshold system is over $\mathbb{R}^N$ (one cutoff per firm). When each firm instead chooses a *subset* of actions—e.g., which countries to produce in—the binary framework cannot capture within-firm complementarities, and the threshold system grows to $\mathbb{R}^M$ (one cutoff per player-action pair). I extend their result to these combinatorial action spaces, where within-firm actions interact through the CDCP's supermodularity and payoffs depend on an endogenous pricing equilibrium. The two results are complementary: Espin-Sanchez et al. cover general payoff functions with binary actions, while I handle combinatorial action spaces under CES/AB payoffs; neither nests the other as a special case. Three results sharpen the distinction. First, a $2 \times 2$ worked example (Section 3.3) constructs a configuration where the binary-action threshold system of Espin-Sanchez et al. cannot be applied, but the combinatorial CDCP structure resolves uniqueness by internalizing within-firm complementarities. Second, the pricing feedback reaches $\rho(\Lambda) = 0.574$ at empirically relevant parameterizations (Section 5.5, Experiment 5), so the endogenous-markup channel is quantitatively significant. Third, endogenous AB markups *expand* the uniqueness region relative to constant CES markups (Section 5.5, Experiment 7): markup compression dampens the dominant-firm responses that drive multiplicity under constant markups.

A key negative result motivates the present approach: Doraszelski and Satterthwaite (2010) introduce private scrap values in dynamic Ericson–Pakes models, guaranteeing cutoff equilibria but explicitly demonstrating that multiplicity persists—private information alone does not resolve the number of equilibria. In the present setting, the pricing contraction (Proposition 11) provides the additional structure—bounding how activation changes feed back through markups—that converts the threshold system into a contrac-

tion. An alternative selection mechanism is the global-games approach of Harrison and Jara-Moroni (2021), who extend iterated dominance arguments to binary-action games with strategic substitutes: dominance regions propagate from extreme types inward until a unique strategy profile survives. Their approach requires binary action spaces and strict payoff asymmetry across player pairs, while the present paper handles combinatorial actions and operates through a different mechanism (pricing contraction rather than dominance propagation).

**Supermodular games.** The lattice-theoretic approach connects to the literature on supermodular games (Milgrom and Roberts 1990; Vives 1990; Milgrom and Shannon 1994; Topkis 1998; Amir 1996). The key challenge is that the cross-player interaction exhibits strategic *substitutes*, not complements: when one firm activates more sites, rivals' activation incentives weaken. I handle this through the squared best-response map, whose monotonicity restores the Tarski fixed-point structure. Athey (2001) establishes existence of pure-strategy equilibria in games whose payoffs satisfy a single-crossing condition in own action and type; the within-player complementarity of the CDCP satisfies her conditions, but the across-player substitutes violate single-crossing in rivals' actions, requiring the additional bounding argument of Proposition 8.

**Combinatorial discrete choice.** Jia (2008) studies store location with strategic substitutes; Arkolakis et al. (2023) formalize the general CDCP framework; Antràs et al. (2017) prove supermodularity under constant CES markups when $(\sigma - 1)/\theta > 1$. I extend their supermodularity result to endogenous variable markups, where activating an additional site changes market shares and hence all firms' markups—introducing a feedback channel absent under constant markups. The proof shows that this feedback is bounded by the same $(\sigma - 1)/\theta > 1$ condition (Proposition 6).

**Variable markups and multinational production.** The application draws on the variable markup literature (Atkeson and Burstein 2008; Edmond et al. 2015; Gaubert and Itskhoki 2021), which takes the set of active firms as given, and the multinational production literature (Tintelnot 2017; Antràs et al. 2026; Castro-Vincenzi 2024), which uses constant CES markups. By closing the loop between activation and pricing, I introduce the three-way circularity (activation $\rightarrow$ market shares $\rightarrow$ markups $\rightarrow$ activation incentives) and show

that the pricing contraction bounds this feedback, preserving both supermodularity and equilibrium uniqueness. The static game provides the stage-game structure for dynamic oligopoly models of the type studied by Ericson and Pakes (1995) and Ifrach and Weintraub (2017). In the estimation of such dynamic games, Aguirregabiria and Mira (2007) develop sequential methods that require equilibrium selection at each iteration, while Bajari et al. (2007) propose forward-simulation methods that bypass equilibrium computation altogether; the uniqueness conditions derived here offer a third path, guaranteeing that the static stage game has a single equilibrium to estimate.

## 2. Model

I develop a static model of combinatorial entry with endogenous pricing. Firms choose which production sites to activate across multiple locations and compete in quantities under CES demand with endogenous variable markups. The leading interpretation is multinational production in segmented international markets, though the game-theoretic results apply more broadly. Figure 1 summarizes the three-way circularity and its relationship to the existing literature.

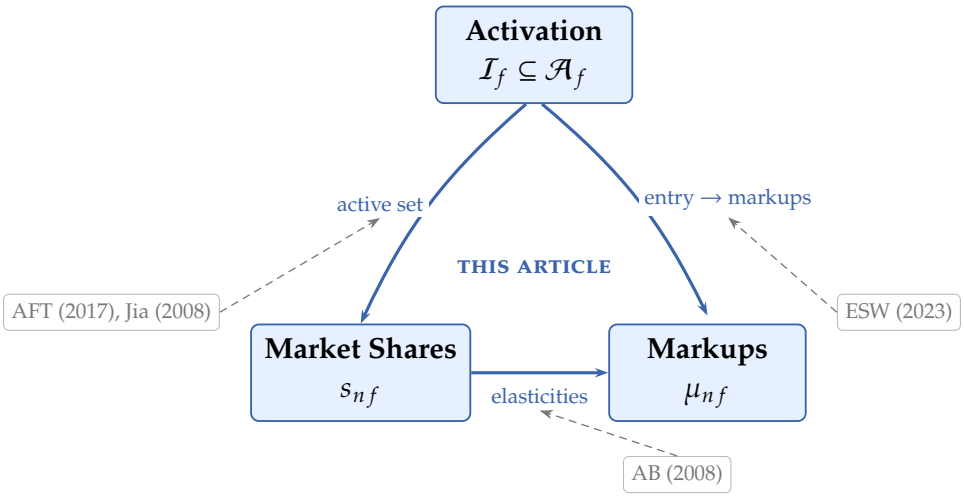


Figure 1: Three-way circularity and the literature. Solid arrows show the full equilibrium loop closed in this article. Prior work addresses subsets: Atkeson and Burstein (2008) endogenize markups given active firms (bottom edge); Antràs et al. (2017) and Jia (2008) solve activation under constant markups (left edge); Espin-Sanchez et al. (2023) prove uniqueness for binary entry decisions given a fixed payoff function (right edge; their result covers the activation channel but does not close the pricing loop).

## 2.1 Environment

There are $N$ countries, indexed by $n \in \{1, \ldots, N\}$. Each country serves as both a potential production site and a destination market. Wages $w_n$ are exogenous.

There are $F$ firms, indexed by $f \in \mathcal{F} = \{1, \ldots, F\}$, that produce differentiated varieties in a single sector. Each firm $f$ is characterized by two primitives:

(i) a *total factor productivity* (TFP) $a_f > 0$, and

(ii) a *set of available production sites* $\mathcal{A}_f \subseteq \{1, \ldots, N\}$.

The firm's strategic decision is which sites in $\mathcal{A}_f$ to activate for production. Section 3.2 extends the model by introducing private information over site-specific fixed costs, yielding a Bayesian game.

## 2.2 Demand

A representative household in each country has income $E_n > 0$ and nested CES preferences. Following Atkeson and Burstein (2008), there is a continuum of sectors indexed by $k \in [0, 1]$, aggregated with elasticity of substitution $\eta > 1$:

$$C_n = \left[ \int_0^1 C_{nk}^{\frac{\eta-1}{\eta}} \, dk \right]^{\frac{\eta}{\eta-1}}. \tag{2.1}$$

The $F$ oligopolistic firms compete in a single differentiated-good sector (say $k = 0$). Because this sector has measure zero in the continuum, its price index $P_n$ does not affect the aggregate price index $\tilde{P}_n$, so the demand shifter $\Omega_n$ defined below is truly exogenous to the firms' decisions. Within the differentiated sector, varieties are aggregated through a CES technology with elasticity $\sigma > \eta$:

$$Q_n = \left[ \sum_{f \in \mathcal{F}_n} q_{nf}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{2.2}$$

where $\mathcal{F}_n$ is the set of firms actively supplying market $n$. Since all production sites can serve all markets via iceberg trade costs, $f \in \mathcal{F}_n$ if and only if $\mathcal{I}_f \neq \emptyset$: a firm is present in

every market once it activates at least one site. The within-sector price index is:

$$P_n = \left[ \sum_{f \in \mathcal{F}_n} p_{nf}^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{2.3}$$

Two-stage budgeting implies that total expenditure on the differentiated sector in market $n$ is:

$$R_n = \Omega_n \, P_n^{1-\eta}, \tag{2.4}$$

where $\Omega_n \equiv E_n / \tilde{P}_n^{1-\eta} > 0$ collects the household's income and the aggregate price index $\tilde{P}_n$. Because the differentiated sector is measure zero in the continuum, $\tilde{P}_n$ is determined by the remaining sectors and is therefore exogenous to the oligopolistic firms. Firm $f$'s demand in market $n$ is:

$$q_{nf} = p_{nf}^{-\sigma} \, P_n^{\sigma-\eta} \, \Omega_n. \tag{2.5}$$

*Remark* 1 (Fixed versus endogenous sectoral expenditure). In the numerical experiments of Section 5, I treat $R_n$ as fixed at its initial value, suppressing the price-index feedback $P_n \to R_n$. This is expected to be conservative for uniqueness: the omitted channel ($P_n^{1-\eta}$ amplifying demand responses through $R_n$) would strengthen pricing feedback and tighten the contraction condition, though a formal proof would require bounding the additional $P_n \to R_n$ loop jointly with the pricing contraction. The Jia–Tarski gap of zero at all tested parameterizations computationally certifies that the equilibrium under fixed $R_n$ satisfies supermodularity. Extending to fully endogenous $R_n$ would add a sector-level general-equilibrium loop; the pricing contraction argument of Proposition 11 would need to absorb the additional $P_n \to R_n \to D_{-f}$ channel, which does not change the structure of the contraction condition.

## 2.3 Production Technology and Costs

Firms produce using labor with heterogeneous productivity across production sites. Firm $f$ with TFP $a_f$ delivering one unit of output from production site $i$ to market $n$ faces unit cost:

$$c_{fin}(\omega) = \frac{w_i \, \tau_{in}}{a_f \, z(\omega)}, \tag{2.6}$$

where $w_i$ is the wage at site $i$, $\tau_{in} \geq 1$ is the iceberg trade cost from site $i$ to market $n$, and $z(\omega)$ is an i.i.d. Fréchet productivity draw with shape parameter $\theta > 0$ across varieties $\omega \in [0, 1]$.

A firm producing at multiple sites sources each variety from the lowest-cost site. As shown by Eaton and Kortum (2002)—and applied at the firm level by Tintelnot (2017) and Antràs et al. (2017)—the Fréchet assumption implies that the firm-level cost index aggregates across active production sites:

$$c_{nf} = \frac{\Gamma_\theta}{a_f} \left[ \sum_{i \in \mathcal{I}_f} \left( w_i \, \tau_{in} \right)^{-\theta} \right]^{-\frac{1}{\theta}}, \tag{2.7}$$

where $\mathcal{I}_f \subseteq \mathcal{A}_f$ is the set of production sites that firm $f$ activates and $\Gamma_\theta = \Gamma\left(1 + \frac{1-\sigma}{\theta}\right)^{\frac{1}{1-\sigma}}$ with $\Gamma(\cdot)$ the Gamma function. Higher TFP $a_f$ lowers the cost index uniformly across all markets. Because $\Gamma_\theta$ is common across all firms, I normalize $\Gamma_\theta = 1$ without loss of generality; it can be absorbed into the TFP parameters $\{a_f\}$.

*Remark* 2 (Cost Complementarity). Equation (2.7) embodies a key complementarity: activating an additional production site $i'$ adds a term $(w_{i'} \tau_{i'n})^{-\theta}$ to the sum inside the brackets, *reducing* the cost index $c_{nf}$ in every market $n$. This cost reduction is larger when the firm already has fewer active sites—a form of increasing differences in the activation indicators. Antràs et al. (2017) establish that this cost complementarity generates supermodularity in the global sourcing problem under constant CES markups when $(\sigma - 1)/\theta > 1$.

## 2.4 Pricing and Markups

Firms compete à la Cournot: each firm chooses the quantity $q_{nf}$ of its variety to supply to market $n$, taking rivals' quantities as given. The nested CES demand system with a finite number of firms generates endogenous variable markups through a mechanism first analyzed by Atkeson and Burstein (2008).

Firm $f$'s revenue share in market $n$ is:

$$s_{nf} \equiv \frac{p_{nf}^{1-\sigma}}{P_n^{1-\sigma}} = \frac{p_{nf} \, q_{nf}}{\sum_{f'} p_{nf'} \, q_{nf'}}. \tag{2.8}$$

Under Cournot competition, the firm internalizes its effect on the sectoral aggregate $Q_n$, yielding a *perceived inverse elasticity* of demand (Appendix C):

$$\frac{1}{\epsilon_{nf}} = \frac{1 - s_{nf}}{\sigma} + \frac{s_{nf}}{\eta}. \tag{2.9}$$

The first-order condition $p_{nf}(1 - 1/\epsilon_{nf}) = c_{nf}$ yields the optimal price as a markup over marginal cost:

$$p_{nf} = \mu_{nf} \cdot c_{nf}, \qquad \mu_{nf} \equiv \frac{\epsilon_{nf}}{\epsilon_{nf} - 1}. \tag{2.10}$$

*Remark* 3 (Bounded Markups and the Cannibalization Channel). Equation (2.9) implies that the perceived inverse elasticity is a harmonic-mean weight between the two CES parameters $\sigma$ and $\eta$:

(i) **Small firm** ($s_{nf} \to 0$): $\epsilon_{nf} \to \sigma$, $\mu_{nf} \to \sigma/(\sigma - 1)$. The firm faces the within-sector CES markup.

(ii) **Dominant firm** ($s_{nf} \to 1$): $\epsilon_{nf} \to \eta$, $\mu_{nf} \to \eta/(\eta - 1)$. The firm's pricing power is bounded by across-sector substitution.

Markups are bounded: $\mu_{nf} \in \left[\sigma/(\sigma - 1),\ \eta/(\eta - 1)\right]$. Activating additional production sites lowers marginal cost (2.7), which raises output and market share, thereby *compressing* the perceived elasticity and *raising* the markup. But the markup is bounded above by $\eta/(\eta - 1)$, so the self-cannibalization force is finite. This *bounded* cannibalization—a direct consequence of the nested CES structure (Atkeson and Burstein, 2008)—competes against the cost complementarity of Remark 2. Whether supermodularity of the activation problem survives this tension is the central question of Section 3. A second, cross-firm channel arises because activation changes the rival demand index (2.11), feeding back into all firms' markups—a pricing interaction absent under constant markups that creates a new obstacle to uniqueness (Proposition 11).

*Remark* 4 (Cournot vs. Bertrand). The perceived elasticity formula (2.9) is identical under Bertrand competition with CES demand (Atkeson and Burstein, 2008), because the nested CES structure implies that the Cournot and Bertrand first-order conditions take the same functional form at a given vector of market shares. The pricing contraction analysis (Proposition 11) and the supermodularity result (Proposition 6) therefore carry over to the

Bertrand case. The models differ in equilibrium shares and markup levels—Bertrand typically yields lower markups—which may shift whether the contraction modulus satisfies $\gamma_{\max} < 1$, but the theoretical structure is unchanged. The Cournot assumption is natural for the multinational production application, where firms choose production scale and capacity.

**Rival demand index.** The competitive environment facing firm $f$ in market $n$ is summarized by a *rival demand index*:

$$D_{-f}(n) = \sum_{f' \neq f,\ f' \in \mathcal{F}_n} p_{nf'}^{1-\sigma}, \tag{2.11}$$

which is a sufficient statistic for competition: firm $f$ needs only $D_{-f}(n)$ to compute its optimal price and variable profit in market $n$. The firm's market share can be written as:

$$s_{nf} = \frac{p_{nf}^{1-\sigma}}{p_{nf}^{1-\sigma} + D_{-f}(n)}, \tag{2.12}$$

making explicit the dependence of the markup (2.9) on the rival demand index.

## 2.5 The Activation Problem

Each firm $f$ chooses which production sites to activate. Let $I_{fi} \in \{0, 1\}$ be an indicator equal to 1 if firm $f$ activates production site $i \in \mathcal{A}_f$, and 0 otherwise. The firm's *activation set* is:

$$\mathcal{I}_f = \left\{ i \in \mathcal{A}_f : I_{fi} = 1 \right\}. \tag{2.13}$$

Activating production site $i$ incurs a fixed operating cost $\phi_i \cdot w_i$ in units of local labor, where $\phi_i$ may vary across countries (reflecting, e.g., differences in regulatory or infrastructure costs).

Firm $f$'s total static profit, given its own activation set $\mathcal{I}_f$ and the equilibrium activation sets of all rivals $\mathcal{I}_{-f}^*$, is:

$$\pi_f\left(\mathcal{I}_f;\, \mathcal{I}_{-f}^*\right) = \sum_{n=1}^{N} \left[ \mathrm{rev}_{nf}\left(\mathcal{I}_f,\, \mathcal{I}_{-f}^*\right) \cdot \frac{\mu_{nf} - 1}{\mu_{nf}} \right] - \sum_{i \in \mathcal{I}_f} \phi_i\, w_i, \tag{2.14}$$

where $\text{rev}_{nf} = p_{nf} \cdot q_{nf}$ is firm $f$'s revenue in market $n$ and $(\mu_{nf} - 1)/\mu_{nf}$ is the variable profit margin. Revenue and markups depend on the full industry activation profile: own activation enters through the cost index (2.7), and rivals' activation sets enter through the rival demand index $D_{-f}(n)$ defined in (2.11).

## 2.6 Static Equilibrium

**Definition 5** (Static Equilibrium)**.** A *static equilibrium* is a profile $\left(\{\mathcal{I}_f^*\}_{f\in\mathcal{F}},\ \{\mu_{nf}^*\},\ \{s_{nf}^*\}\right)$ of activation sets, markups, and market shares such that:

**E1. Profit maximization:** For each firm $f$, the activation set $\mathcal{I}_f^*$ solves

$$\mathcal{I}_f^* \in \arg\max_{\mathcal{I}_f \subseteq \mathcal{A}_f} \pi_f\left(\mathcal{I}_f;\, \mathcal{I}_{-f}^*\right). \tag{2.15}$$

**E2. Markup consistency:** For each $(n, f)$ with $\mathcal{I}_f^* \neq \emptyset$, the markup $\mu_{nf}^*$ satisfies the pricing equation (2.10) with perceived elasticity (2.9), and market shares $s_{nf}^*$ are consistent with equilibrium prices via (2.12).

The static equilibrium involves a three-way circularity:

$$\underbrace{\text{Activation sets } \mathcal{I}_f}_{\text{combinatorial choice}} \longrightarrow \underbrace{\text{Market shares } s_{nf}}_{\text{CES aggregation}} \longrightarrow \underbrace{\text{Markups } \mu_{nf}}_{\text{AB pricing}} \longrightarrow \underbrace{\text{Profitability of } \mathcal{I}_f}_{\text{back to activation}}.$$

With constant markups (as in Antràs et al., 2017), the last link is severed: activation does not feed back through markups. The theoretical contribution of this article is to show that the CDCP structure—and in particular, supermodularity—survives this feedback (Section 3). A complementary computational contribution compares alternative algorithms for resolving the three-way fixed point (Section 4).

# 3. Theoretical Results

This section establishes the theoretical properties of the static game. I first show that each firm's activation problem is supermodular (Proposition 6), then characterize the equilibrium set (Proposition 8) and establish convergence (Proposition 9). The main

result—BNE uniqueness via a Jacobian decomposition—builds on the pricing-uniqueness condition of Section 3.1 and is stated formally in Section 3.2.

Since the rival demand index $D_{-f}(n)$ is a sufficient statistic for competition (Section 2.4), I write $\pi_f(\mathcal{I}_f; D_{-f})$ for firm $f$'s profit (2.14) evaluated at the unique pricing equilibrium given $D_{-f}$.

**Proposition 6** (Single-Agent Supermodularity)**.** *Fix the rival demand index $D_{-f}(n)$ for all markets $n$ and suppose $(\sigma-1)/\theta > 1$.*

*(i)* Per-market sufficient condition. *If the equilibrium market shares satisfy $\max_{n,f} s_{nf} < \bar{s}^*$, where $\bar{s}^* \equiv (\alpha-1)/(2\alpha)$ with $\alpha = (\sigma-1)/\theta$ is the CES-benchmark per-market convexity threshold (Appendix A, Step 4), then each market's contribution to variable profit is convex in the cost aggregator, which implies that firm $f$'s activation profit $\pi_f(\mathcal{I}_f; D_{-f})$ exhibits* increasing differences *in the activation indicators $\{I_{fi}\}_{i\in\mathcal{A}_f}$:*

$$\pi_f\big(\mathcal{I}_f \cup \{i,i'\}; D_{-f}\big) - \pi_f\big(\mathcal{I}_f \cup \{i\}; D_{-f}\big) \geq \pi_f\big(\mathcal{I}_f \cup \{i'\}; D_{-f}\big) - \pi_f\big(\mathcal{I}_f; D_{-f}\big). \quad (3.1)$$

*(ii)* Computational certification. *When some market shares exceed $\bar{s}^*$, the per-market condition is not met, but* aggregate *supermodularity across all markets can still hold because low-share markets (positive curvature) dominate high-share markets (negative curvature). The Jia–Tarski algorithm certifies aggregate supermodularity whenever the upper and lower bounds coincide (gap = 0).*

*Under either (i) or (ii), the set of optimal solutions to $\max_{\mathcal{I}_f\subseteq\mathcal{A}_f} \pi_f(\mathcal{I}_f; D_{-f})$ has a greatest element and a least element, each attainable by the Jia–Tarski algorithm.*

*Proof sketch.* The marginal return to activating site $i'$ can be decomposed into two effects. The *cost complementarity effect*: by equation (2.7), adding site $i'$ contributes a term $(w_{i'}\tau_{i'n})^{-\theta}$ to the cost index in every market $n$. The proportional cost reduction is larger when the existing cost index is higher—i.e., when fewer other sites are already active. Under the condition $(\sigma-1)/\theta > 1$, variable profit is sufficiently convex in the cost index that the scale expansion from lower costs generates increasing differences.

The *self-cannibalization effect*: activating $i'$ lowers cost, which lowers price, which raises market share $s_{nf}$, which compresses the perceived elasticity $\epsilon_{nf}$ via (2.9) and

raises the markup $\mu_{nf}$. This partially offsets the quantity gain. Under the CES benchmark (constant markups), per-market convexity of variable profit in the cost aggregator $\Sigma_{fn} \equiv \sum_{i \in \mathcal{I}_f} (w_i \tau_{in})^{-\theta}$ from (2.7) requires $s < \bar{s}^*$; under AB endogenous markups, the pricing damping raises the effective threshold (Appendix A, Steps 3–4). When some shares exceed $\bar{s}^*$, aggregate supermodularity is verified via part (ii).

The condition $(\sigma - 1)/\theta > 1$ is the same as in Antràs et al. (2017), Theorem 1, who establish supermodularity under constant markups. The extension here shows that the additional markup feedback does not overturn the result. Full proof in Appendix A. □

Under Atkeson–Burstein endogenous markups, the pricing response is further damped (Appendix A, Step 3), raising the effective convexity threshold above $\bar{s}^*$. At the baseline parameterization of Section 5, $\max_{n,f} s_{nf} = 0.39 > \bar{s}^* = 0.125$, yet the Jia–Tarski gap is zero at all iterations, certifying aggregate supermodularity via part (ii).

*Remark* 7 (Supermodularity in practice). Propositions 8–13 require only that the single-agent activation problem is supermodular—whether certified analytically via part (i) or computationally via part (ii) is immaterial to the argument. Note that the per-market condition in part (i) must hold across *all* possible activation sets $\mathcal{I}_f$, not only the equilibrium one, because the market share $s_{nf}$ depends on the cost index $\Sigma_{fn}(\mathcal{I}_f)$ which varies with the activation set. The computational certification in part (ii) verifies supermodularity at all relevant activation sets via the Jia–Tarski algorithm and is the recommended diagnostic.

**Existence.** A Nash equilibrium of the complete-information game exists by Nash's theorem, since the activation lattice $L = \prod_f 2^{\mathcal{A}_f}$ is finite. The following two propositions characterize the equilibrium set and establish convergence of iterative algorithms.

**Proposition 8** (Equilibrium Bounds)**.** *Suppose each firm's activation profit is supermodular (Proposition 6).*

*(i) The joint best-response map $\Phi$ is* anti-monotone *(order-reversing) on the joint activation lattice $L$, so the squared map $\Phi^2 = \Phi \circ \Phi$ is monotone. By Tarski's (1955) fixed-point theorem, $\Phi^2$ has a greatest fixed point $\bar{\mathbf{I}}$ and a least fixed point $\underline{\mathbf{I}}$ on $L$.*

*(ii) Any Nash equilibrium $\mathbf{I}^*$ satisfies $\underline{\mathbf{I}} \leq \mathbf{I}^* \leq \bar{\mathbf{I}}$: the extremal fixed points of $\Phi^2$ bound the entire Nash equilibrium set.*

*(iii) If $\bar{\mathbf{I}}$ is itself a Nash equilibrium (i.e., $\Phi(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$, not merely a two-cycle of $\Phi$), it is the greatest Nash equilibrium.*

*Proof.* Since activation choices are *strategic substitutes*—when rival firms activate more sites, the rival demand index $D_{-f}$ increases, reducing the profitability of firm $f$'s sites—the joint best-response map $\Phi : (\mathcal{I}_1, \ldots, \mathcal{I}_F) \mapsto (\mathcal{I}_1^*, \ldots, \mathcal{I}_F^*)$ is *anti-monotone*: $\mathbf{I} \leq \mathbf{I}'$ implies $\Phi(\mathbf{I}) \geq \Phi(\mathbf{I}')$ (cf. Amir, 1996), where $\leq$ is the component-wise set-inclusion order on $L \equiv \prod_f 2^{\mathcal{A}_f}$.

Part (i): $\Phi^2$ is monotone because anti-monotonicity composes to give $\mathbf{I} \leq \mathbf{I}' \Rightarrow \Phi^2(\mathbf{I}) \leq \Phi^2(\mathbf{I}')$. Tarski's theorem on the finite complete lattice $L$ yields the extremal fixed points.

Part (ii): any Nash equilibrium satisfies $\Phi(\mathbf{I}^*) = \mathbf{I}^*$, hence $\Phi^2(\mathbf{I}^*) = \mathbf{I}^*$. By Tarski's theorem, $\underline{\mathbf{I}} \leq \mathbf{I}^* \leq \bar{\mathbf{I}}$.

Part (iii): if $\Phi(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$, then $\bar{\mathbf{I}}$ is a Nash equilibrium; any other Nash equilibrium $\mathbf{I}^*$ satisfies $\mathbf{I}^* \leq \bar{\mathbf{I}}$ by (ii).

In general, $\bar{\mathbf{I}}$ may be a two-cycle rather than a Nash equilibrium: it is possible that $\Phi(\bar{\mathbf{I}}) \neq \bar{\mathbf{I}}$ while $\Phi^2(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$. However, at all parameterizations tested in Section 5, $\Phi(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$ and the equilibrium is unique. Full proof in Appendix B. □

**Proposition 9** (Convergence of Jacobi Iteration)**.** *Suppose each firm's activation profit is supermodular. Starting from the supremum $\hat{\mathbf{I}} = (\mathcal{A}_1, \ldots, \mathcal{A}_F)$, the even subsequence $\{\Phi^{2k}(\hat{\mathbf{I}})\}$ is decreasing and converges to $\bar{\mathbf{I}}$ (the greatest fixed point of $\Phi^2$). The odd subsequence $\{\Phi^{2k+1}(\hat{\mathbf{I}})\}$ is increasing and converges to $\Phi(\bar{\mathbf{I}})$. If $\Phi(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$, the full sequence converges to the greatest Nash equilibrium. Analogously from the infimum.*

*Proof.* Since $\Phi$ is anti-monotone and $\Phi(\hat{\mathbf{I}}) \leq \hat{\mathbf{I}}$, the even subsequence $\{\Phi^{2k}(\hat{\mathbf{I}})\}$ is decreasing (by monotonicity of $\Phi^2$) and the odd subsequence $\{\Phi^{2k+1}(\hat{\mathbf{I}})\}$ is increasing, with $\Phi^{2k+1}(\hat{\mathbf{I}}) \leq \Phi^{2k}(\hat{\mathbf{I}})$ for all $k$. On the finite lattice $L$, both subsequences stabilize in finitely many steps. The even limit is $\bar{\mathbf{I}}$ and the odd limit is $\Phi(\bar{\mathbf{I}})$, which satisfy $\Phi^2(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$ and $\Phi(\bar{\mathbf{I}}) \leq \bar{\mathbf{I}}$. Full proof in Appendix B. □

*Remark* 10 (Gauss–Seidel convergence). The sequential update map $\Psi$ that updates firms in order $(f_1, \ldots, f_F)$ using the latest available rivals' strategies converges at all tested parameterizations (Section 5). Formal convergence of $\Psi$ under anti-monotone best responses is not guaranteed in general—sequential updating breaks the even/odd subsequence

structure that yields monotonicity of $\Phi^2$—but empirical convergence is robust across all parameterizations tested.

At all parameterizations tested in Section 5, $\Phi(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$: both Jacobi and Gauss–Seidel converge to the same Nash equilibrium, and no two-cycles are observed. Gauss–Seidel uses updated rival information within each round and typically converges in fewer rounds, whereas Jacobi is naturally parallelizable. Section 4 presents the full comparison.

## 3.1 Pricing Uniqueness and the Intensive Margin

The uniqueness question has two distinct margins. On the *intensive margin* (pricing), uniqueness requires that the AB markup fixed point is contractive for any given activation profile. On the *extensive margin* (site activation), uniqueness requires that no "marginal sites"—sites whose profitability depends on rival behavior—switch activation status across equilibria. This subsection resolves the intensive margin via a contraction condition on the feedback coefficient (Proposition 11); Section 3.2 resolves the extensive margin by introducing private information over fixed costs.

**Feedback coefficient.** The strength of the strategic interaction through the AB markup channel is captured by a *feedback coefficient*. For firm $f$ in market $n$ with share $s_{nf}$ and markup $\mu_{nf}$, define:

$$\kappa_{nf} = \frac{s_{nf}^2 \, \beta_{nf}}{1 + s_{nf}(1 - s_{nf}) \, \beta_{nf}}, \qquad \beta_{nf} = \frac{(\sigma - 1) \, \mu_{nf} \, (\sigma - \eta)}{\sigma \, \eta}. \tag{3.2}$$

The coefficient $\kappa_{nf} = \partial p_{nf}^{1-\sigma} / \partial D_{-f}(n)$ measures how firm $f$'s pricing power responds to a change in the rival demand index, holding the activation set fixed. When $s_{nf} = 0$ (negligible firm), $\kappa_{nf} = 0$; as $s_{nf}$ rises, the markup channel amplifies the pricing response, so $\kappa_{nf}$ increases.

**Proposition 11** (Pricing Uniqueness)**.** *If*

$$\max_{n,f} \kappa_{nf} \; < \; \frac{1}{F - 1}, \tag{3.3}$$

*where $\kappa_{nf}$ is the feedback coefficient (3.2), then for any fixed activation profile $(\mathcal{I}_1, \ldots, \mathcal{I}_F)$, the AB*

*pricing equilibrium is unique.*[1]

*Proof.* For a fixed activation profile, the AB pricing system defines a map $\Psi : D \mapsto D'$ on the space of rival demand indices. Each entry of the Jacobian of $\Psi$ is bounded by $\kappa_{nf}$. Under (3.3), the row sums of the interaction matrix satisfy $\sum_{g \neq f} |\partial D'_{-f}(n)/\partial D_{-g}(n)| \leq (F-1)\max \kappa < 1$, so $\Psi$ is a contraction under the sup-norm. By the Banach fixed-point theorem, the pricing equilibrium is unique for every activation profile. □

The condition (3.3) uses the infinity-norm bound and may be conservative. A tighter diagnostic uses the spectral radius of the interaction matrix:

$$\Lambda_{(n,f),(n,g)} \;=\; \kappa_{nf} \cdot \frac{p_{ng}^{1-\sigma}}{D_{-f}(n)}, \qquad g \neq f, \tag{3.4}$$

with all other entries zero. The matrix $\Lambda \in \mathbb{R}^{NF \times NF}$ is block-diagonal across markets ($N$ blocks of size $F \times F$), so $\rho(\Lambda) = \max_n \rho(\Lambda_n)$ is cheap to compute, and pricing uniqueness holds whenever $\rho(\Lambda) < 1$.

Proposition 11 resolves the intensive margin but does *not* guarantee uniqueness of the full game, because different activation profiles lead to different demand indices. Define the composite map $\tilde{\Phi} : D \mapsto D'$ that maps rival demands to best-response activation sets, then evaluates the unique pricing equilibrium to produce new demands. Multiple equilibria arise when $\tilde{\Phi}$ has multiple fixed points—equivalently, when there exist marginal sites whose activation profit $\Delta\pi_{if}$ changes sign across equilibrium demands. The number of marginal sites is the key diagnostic for multiplicity; Section 5.5 shows it is zero at empirically relevant parameterizations. To resolve the extensive margin analytically, the next subsection introduces private information over fixed costs, converting the discrete activation problem into a continuous threshold system on which contraction arguments apply.

[1] Condition (3.3) is sufficient but conservative. At the baseline parameterization of Section 5, $\max_{n,f} \kappa_{nf} = 0.193 > 1/(F-1) = 0.143$, so (3.3) is violated. Pricing uniqueness is instead confirmed by the tighter spectral condition $\rho(\Lambda) = 0.096 \ll 1$ (equation 3.4). For applied work, the spectral radius is the recommended diagnostic.

## 3.2 Bayesian Extension with Private Costs

In practice, firms' fixed costs are not publicly observed. I model this by introducing private information over fixed costs, yielding a Bayesian game whose equilibrium is characterized by a threshold system on which contraction-based uniqueness arguments can be applied.

**Assumption 12** (Private cost shocks)**.** Each firm $f$ privately observes fixed-cost shocks $\{\varepsilon_{if}\}_{i\in\mathcal{A}_f}$ that are i.i.d. across sites and across firms, drawn from a distribution $G$ with full-support density $g$, mean zero, and scale parameter $\sigma_\varepsilon > 0$. The total fixed cost of activating site $i$ is $(\phi_i + \varepsilon_{if})\, w_i$. Firm $f$ observes its own draws but not rivals'.

The i.i.d. assumption is substantive: it ensures that each site's activation decision reduces to a scalar threshold comparison (Proposition 13(i)), enabling the Jacobian factorization (3.6). With correlated shocks—for example, a hierarchical structure $\varepsilon_{if} = \eta_f + \tilde{\varepsilon}_{if}$ with a common firm-level component $\eta_f$—the activation decision for site $i$ would depend on $\eta_f$ through the conditional expected marginal value, yielding a threshold that varies with the firm-level draw rather than an unconditional scalar. The Jacobian factorization breaks because the density factor becomes conditional on $\eta_f$. A partial extension is possible: conditional on $\eta_f$, the site-level shocks $\tilde{\varepsilon}_{if}$ remain i.i.d., so a conditional version of the contraction argument applies at each realization, but the fixed point is then over a higher-dimensional system (one threshold per site per firm-level type) and the contraction condition is harder to verify. In practice, the i.i.d. assumption is standard in the CDCP literature (Jia, 2008; Arkolakis et al., 2023); the restriction here is no more binding than in those frameworks.

Economically, the i.i.d. assumption rules out two forms of dependence that arise naturally in multinational production. First, *within-firm correlation*: a firm-level "headquarters quality" draw that shifts fixed costs at all sites simultaneously—for example, a management practice that raises or lowers setup costs uniformly across countries. Second, *spatial correlation*: regulatory or institutional costs that co-move across neighboring countries sharing a border or regulatory regime. Either form of dependence would break the one-dimensional threshold structure exploited in Proposition 13(i).

Two distinct roles of the i.i.d. assumption deserve emphasis. Independence across sites *within* a given firm is standard in the CDCP literature: it ensures that each site's activation reduces to a scalar threshold comparison, enabling the Jia–Tarski algorithm.

Independence *across firms* does additional work here: it makes each firm's threshold an unconditional scalar that does not depend on other firms' cost realizations, so the cross-player Jacobian $J_T$ has the product structure (3.6) that supports the contraction argument. With correlated shocks across firms—for example, a common macroeconomic shock—the thresholds would be conditional on the common state, and the contraction condition would need to hold state-by-state.

The solution concept is Bayesian Nash equilibrium (BNE). Existence is immediate: the threshold map $T : [0,1]^M \to [0,1]^M$ (defined below) is continuous, so Brouwer's fixed-point theorem guarantees a BNE in threshold strategies.

**Threshold strategies.** The single-agent activation problem is supermodular in $(\mathcal{I}_f, -\varepsilon_f)$: higher fixed-cost shocks make activation less attractive, and the interaction is complementary by Proposition 6. By the monotone comparative statics of Athey (2001), the BNE involves *threshold strategies*: firm $f$ activates site $i$ if and only if $\varepsilon_{if} < \bar{\varepsilon}_{if}$, where $\bar{\varepsilon}_{if}$ is an endogenous threshold. Given rivals' thresholds, each site $i$ of rival $g$ is active independently with probability $\nu_{ig} = G(\bar{\varepsilon}_{ig})$.

**Threshold system.** The threshold $\bar{\varepsilon}_{if}$ is pinned down by the zero-marginal-value condition: at $\varepsilon_{if} = \bar{\varepsilon}_{if}$, the expected marginal value of activating site $i$, integrating over rivals' random activation sets, equals zero. Define the best-response map $T$ that maps rival activation probabilities $\{\nu_{ig}\}_{g \neq f}$ to own activation probabilities $\{\nu_{if}\}$ via:

$$\{\nu_{ig}\}_{g\neq f} \xrightarrow{\text{expected demand}} \mathbb{E}[D_{-f}] \xrightarrow{\rho<1} \text{pricing eq.} \xrightarrow{\text{bounded markups}} \bar{\varepsilon}_{if} \xrightarrow{G} \nu_{if}. \quad (3.5)$$

The nested CES demand structure provides three ingredients for this map to be contractive:

1. *Bounded markups* $\mu \in [\sigma/(\sigma-1),\ \eta/(\eta-1)]$: the marginal value of each site is Lipschitz-continuous in the rival demand index, so small changes in the pricing equilibrium produce bounded changes in the threshold $\bar{\varepsilon}_{if}$, with Lipschitz constant depending on $(\sigma, \eta)$.

2. *Pricing contraction* $\rho(\Lambda) < 1$: the map from expected rival demands to the pricing equilibrium is contractive (Proposition 11), ensuring that a small shift in activation probabilities feeds back into prices with bounded magnification.

3. *Smoothing via G*: the map from thresholds to probabilities (via the CDF) is smooth and bounded, converting the discrete extensive margin into a continuous probability.

**Jacobian decomposition.** Before stating the uniqueness result, I develop its key tool: a quantitative bound on the contraction modulus of the threshold map $T$ defined in (3.5). Write $G(x) = G_0(x/\sigma_\varepsilon)$ with standardized CDF $G_0$ and density $g_0 = G_0'$, so that $g(x) = g_0(x/\sigma_\varepsilon)/\sigma_\varepsilon$. The Jacobian of $T$ factors as

$$J_T(v) = \underbrace{\operatorname{diag}\left(\frac{g_0(\bar{\varepsilon}_{if}/\sigma_\varepsilon)}{\sigma_\varepsilon}\right)}_{\text{density dampening}} \cdot \underbrace{\operatorname{diag}\left(\frac{1}{w_i}\right)}_{\text{scaling}} \cdot \underbrace{\left[\frac{\partial \mathbb{E}[\Delta\pi_{if}]}{\partial v_{jg}}\right]}_{\equiv J_\Delta\text{: marginal-value sensitivity}}, \tag{3.6}$$

where $\bar{\varepsilon}_{if} = \mathbb{E}[\Delta\pi_{if}]/w_i$ is the threshold of site $(i, f)$. Define $J_v \equiv \operatorname{diag}(1/w_i) \cdot J_\Delta$, which absorbs the wage scaling into the sensitivity matrix. The three ingredients above appear as three factors governing the contraction modulus $\gamma(\sigma_\varepsilon) \equiv \sup_v \|J_T(v)\|_\infty$:

(a) *Density dampening*: the prefactor $g_0(\cdot)/\sigma_\varepsilon$ vanishes as $\sigma_\varepsilon \to \infty$.

(b) *Marginal-value sensitivity*: the matrix $J_\Delta$ inherits bounded entries from the CES demand structure. The across-firm entries are non-positive (strategic substitutes through $D_{-f}(n)$), while within-firm entries are non-negative (complementarity from Proposition 6). Both are bounded by the markup range $[\sigma/(\sigma-1),\ \eta/(\eta-1)]$, and the pricing contraction amplifies demand perturbations by at most $1/(1 - \|\Lambda\|_\infty)$ (Appendix D).

(c) *Threshold location*: through $g_0(\bar{\varepsilon}_{if}/\sigma_\varepsilon)$, the density is small when the threshold is far from the mode—i.e., when site $i$ is far from marginal.

**Proposition 13** (BNE Uniqueness)**.** *Suppose each firm's activation profit is supermodular and the pricing equilibrium is unique ($\rho(\Lambda) < 1$; a sufficient condition is $\|\Lambda\|_\infty < 1$, guaranteed by Proposition 11).*

*(i) The BNE is in threshold strategies: firm $f$ activates site $i$ iff $\varepsilon_{if} < \bar{\varepsilon}_{if}$.*

*(ii) Define the* contraction modulus

$$\gamma(\sigma_\varepsilon) \equiv \sup_{v \in [0,1]^M} \|J_T(v)\|_\infty \le \frac{\|g_0\|_\infty}{\sigma_\varepsilon} \cdot \|J_v\|_\infty, \tag{3.7}$$

*where* $M = \sum_f |\mathcal{A}_f|$, $\|g_0\|_\infty = \sup_z g_0(z)$ *is the peak density of the standardized distribution, and* $\|J_v\|_\infty = \sup_\nu \max_{(i,f)}(1/w_i) \sum_{(j,g)} |\partial \mathbb{E}[\Delta\pi_{if}]/\partial v_{jg}|$ *is the weighted row-sum norm of the marginal-value sensitivity. If* $\gamma(\sigma_\varepsilon) < 1$*, the threshold map* $T$ *is a contraction and the BNE is unique.*

*(iii) As* $\sigma_\varepsilon \to 0$*, the BNE activation probabilities converge to* $\{0, 1\}$*, recovering a Nash equilibrium of the complete-information game.*

*(iv) Assume additionally that* $G_0$ *has log-concave tails (e.g., normal or logistic). The map* $\sigma_\varepsilon \mapsto \gamma(\sigma_\varepsilon)$ *vanishes at both extremes:* $\gamma \to 0$ *as* $\sigma_\varepsilon \to \infty$ *(density dampening dominates) and* $\gamma \to 0$ *as* $\sigma_\varepsilon \to 0$ *when all sites are non-marginal (thresholds pushed to density tails). If the complete-information game has a unique Nash equilibrium with gap* $\delta_v = \min_{i,f} |\Delta\pi^*_{if}|/w_i > 0$ *and* $\gamma_{\max} \equiv \max_{\sigma_\varepsilon > 0} \gamma(\sigma_\varepsilon) < 1$*, then the BNE is unique for* all $\sigma_\varepsilon > 0$*. The condition* $\gamma_{\max} < 1$ *is verifiable numerically by evaluating* $\gamma$ *on a grid of* $\sigma_\varepsilon$ *values.*

*Proof sketch.* Part (i): conditional on rivals' threshold strategies (which determine activation probabilities $\{v_{jg}\}_{g \neq f}$), firm $f$ faces a single-agent CDCP with stochastic fixed costs $(\phi_i + \varepsilon_{if})w_i$. By Proposition 6, this CDCP is supermodular in the activation indicators. Since the shocks $\varepsilon_{if}$ are i.i.d. across sites, the optimal activation set is increasing in $-\varepsilon_f$ componentwise on the product lattice $\{0, 1\}^{|\mathcal{A}_f|}$ (Topkis 1998, Theorem 4.2.2). It follows that each site $i$ has a threshold $\bar{\varepsilon}_{if}$ (depending on the equilibrium probabilities of all other sites) such that site $i$ is activated iff $\varepsilon_{if} < \bar{\varepsilon}_{if}$. Crucially, $\bar{\varepsilon}_{if}$ is an *unconditional scalar*: it does not depend on firm $f$'s own-site shocks $\varepsilon_{jf}$ for $j \neq i$. This follows directly from Assumption 12: since $\varepsilon_{if} \perp \varepsilon_{jf}$, the conditional expectation $\mathbb{E}[\Delta\pi_{if} \mid \varepsilon_{if}]$ does not depend on other sites' shock realizations, so the activation decision for site $i$ reduces to a one-dimensional comparison.

For (ii), the factorization (3.6) decomposes each row of $|J_T|$ into density dampening ($\leq \|g_0\|_\infty/\sigma_\varepsilon$) times the weighted marginal-value sensitivity $J_v$. Since the expected marginal value is a polynomial in $\{v_{ig}\}$, it is smooth; its row sums are bounded by composing Lipschitz constants along the chain $\nu \to \mathbb{E}[D_{-f}] \to$ pricing $\to \Delta\pi_{if}$, with the pricing contraction amplifying perturbations by at most $1/(1 - \rho(\Lambda))$. This yields $\|J_v\|_\infty \leq L_D \cdot L_\mu/(1 - \|\Lambda\|_\infty) + L_C < \infty$ (Appendix D), so $\gamma \leq (\|g_0\|_\infty/\sigma_\varepsilon) \cdot \|J_v\|_\infty < 1$ for $\sigma_\varepsilon$ large enough; Banach's theorem yields uniqueness.

Part (iii) follows from upper hemicontinuity of the equilibrium correspondence: as $\sigma_\varepsilon \to 0$, the perturbed game converges to the complete-information game, and any sequence of BNE has a limit point that is a Nash equilibrium of the complete-information game (cf. Harsanyi 1973; for entry games with private information, see Espin-Sanchez et al. 2023, Proposition 2).

For (iv), $\gamma \to 0$ as $\sigma_\varepsilon \to \infty$ is immediate from (3.7). For $\sigma_\varepsilon \to 0$: when the complete-information game has a unique equilibrium with gap $\delta_v > 0$, all thresholds are pushed into the density tails, yielding the convergence rate $\gamma(\sigma_\varepsilon) \leq (C_g/\sigma_\varepsilon) \exp(-c\, \delta_v\, w_{\min}/(2\sigma_\varepsilon))\, \|J_v\|_\infty$ for log-concave $g_0$ with tail bound $g_0(z) \leq C_g e^{-c|z|}$ (Appendix D, Step 4). At any $\sigma_\varepsilon$ where $\gamma < 1$, the implicit function theorem ensures that the BNE extends as a unique $C^1$ branch. Since $\gamma \to 0$ at both extremes and $\gamma_{\max} < 1$ implies $\gamma < 1$ everywhere, uniqueness holds for all $\sigma_\varepsilon > 0$. □

The interaction matrix $\Lambda$ controls *both* margins through a single channel. On the intensive margin, $\rho(\Lambda) < 1$ guarantees pricing uniqueness (Proposition 11). On the extensive margin, $\rho(\Lambda)$ enters the contraction modulus (3.7) through the amplification factor $1/(1 - \rho(\Lambda))$: the sensitivity of the pricing equilibrium to a demand perturbation is bounded by $\|(I - \Lambda)^{-1}\| \leq 1/(1 - \rho(\Lambda))$. The infinity-norm condition $\|\Lambda\|_\infty < 1$ of Proposition 11 is a conservative sufficient condition; the spectral condition $\rho(\Lambda) < 1$ is tighter and is used in all numerical diagnostics. At the baseline parameterization of Section 5, $\rho(\Lambda) = 0.096$ (amplification factor 1.11): the outer activation problem is *nearly* a constant-markup entry game. Across the robustness experiments, $\rho(\Lambda)$ ranges from 0.025 (low $\sigma$, high $\eta$) to 0.574 (high $\sigma$, low $\eta$, concentrated market; Section 5.5), remaining well below one throughout. When $\rho(\Lambda) \to 1$, the amplification diverges and larger $\sigma_\varepsilon$ is required to restore the contraction.

*Remark* 14 (Structural vs. CES-specific elements). The Jacobian decomposition (3.6) and the bathtub shape of $\gamma(\sigma_\varepsilon)$ are *structural*: they hold for any Bayesian game with private cost shocks, a contractive pricing map, and Lipschitz-continuous marginal values—regardless of the demand system. What the CES structure provides is closed-form expressions for the Lipschitz constants ($L_D$, $L_\mu$, $L_C$), the feedback coefficient $\kappa_{nf}$, and the share thresholds $\bar{s}^*$ and $\bar{s}$. Under more general demand systems (e.g., translog), markups need not be bounded, so the Lipschitz constants may diverge and the contraction modulus may exceed one. The qualitative insight—that uniqueness is governed by a bathtub-shaped modulus

driven by three interpretable forces—extends beyond CES, but the quantitative verification $\gamma_{\max} < 1$ would require numerical Jacobian bounds tailored to the specific demand system.

*Remark* 15 (Role of private information). Proposition 13 establishes uniqueness of the Bayesian game for a given noise distribution $G$ and scale $\sigma_\varepsilon$. The BNE is a well-defined equilibrium concept in its own right—not merely a device for selecting among complete-information equilibria. Non-marginal sites ($|\Delta\pi_{if}|$ bounded away from zero) activate with probability near one or zero; only marginal sites ($|\Delta\pi_{if}| \approx 0$) have interior activation probabilities that depend on $G$. When the complete-information game has a unique Nash equilibrium (as at all tested parameterizations), the BNE converges to it as $\sigma_\varepsilon \to 0$ regardless of $G$.

The following theorem consolidates all assumptions for the main result in one place, referencing Propositions 6–13 as building blocks.

**Theorem 16** (Main Result: Global BNE Uniqueness)**.** *Suppose:*

*(a)* $(\sigma - 1)/\theta > 1$ *and each firm's activation profit is supermodular (Proposition 6, certified analytically or computationally);*

*(b)* $\rho(\Lambda) < 1$*, where* $\Lambda$ *is the markup interaction matrix (Proposition 11);*

*(c) the standardized distribution* $G_0$ *has log-concave tails (e.g., normal or logistic);*

*(d) the complete-information game has a unique Nash equilibrium with gap* $\delta_v = \min_{i,f} |\Delta\pi^*_{if}|/w_i > 0$*;*

*(e)* $\gamma_{\max} \equiv \max_{\sigma_\varepsilon > 0} \gamma(\sigma_\varepsilon) < 1$.

*Then the Bayesian Nash equilibrium is unique for all* $\sigma_\varepsilon > 0$.

*Proof.* Conditions (a)–(b) ensure that the single-agent CDCP is supermodular and the pricing equilibrium is unique, so the threshold map $T$ in (3.5) is well-defined. Conditions (c)–(e) verify the hypotheses of Proposition 13(iv): the contraction modulus $\gamma(\sigma_\varepsilon)$ vanishes at both extremes and its peak is below one, so $T$ is a contraction for all $\sigma_\varepsilon > 0$. Banach's fixed-point theorem yields a unique fixed point. □

**Sharp market-share threshold.** Theorem 16 consolidates the assumptions; the following corollary provides a closed-form diagnostic. The nested CES demand structure yields a closed-form characterization of the contraction condition. Since $\kappa_{nf}$ depends only on the market share $s_{nf}$ and the elasticity pair $(\sigma, \eta)$, condition (3.3) reduces to a threshold on the maximum market share.

**Corollary 17** (Market-Share Threshold for BNE Uniqueness)**.** *Define $\bar{s}(\sigma, \eta, F)$ as the unique solution in $(0, 1)$ of $\kappa(\bar{s}; \sigma, \eta) = 1/(F-1)$. If $\max_{n,f} s_{nf} < \bar{s}$, then the BNE is unique. For $(\sigma, \eta) = (5, 2)$, the threshold solves*

$$24\bar{s}^2 - 9\bar{s} - 8 = 0, \qquad \bar{s} \approx 0.795 \quad (F = 2). \tag{3.8}$$

*For $F = 8$ firms, $\bar{s} \approx 0.339$; for $F = 4$, $\bar{s} \approx 0.513$.*

*Proof.* From (3.2), for $(\sigma, \eta) = (5, 2)$: $1/\epsilon = (1-s)/5 + s/2 = (2+3s)/10$, so $\mu = 10/(8-3s)$ and $\beta = 12/(8-3s)$. Substituting into $\kappa$:

$$\kappa(s) \;=\; \frac{12s^2}{8 + 9s - 12s^2}.$$

Setting $\kappa = 1$ (i.e., $F = 2$) gives $24s^2 - 9s - 8 = 0$, whose positive root is $\bar{s} = (9 + \sqrt{849})/48 \approx 0.795$. □

Comparative statics follow from $\kappa$ being increasing in $s$ and in $\sigma - \eta$:

1. $\partial\bar{s}/\partial\sigma < 0$: higher within-sector substitution (larger $\sigma - \eta$) strengthens the markup channel, *tightening* the threshold.

2. $\partial\bar{s}/\partial\eta > 0$ (for $\eta < \sigma$): as $\eta \to \sigma$, nested CES collapses to standard CES with constant markups, so $\kappa \to 0$ and $\bar{s} \to 1$.

Table 1 reports $\bar{s}$ for selected $(\sigma, \eta, F)$ combinations, showing that the threshold is tightest when the markup channel is strongest (low $\eta$, high $\sigma - \eta$, large $F$).

*Remark* 18 (Two share thresholds). The paper employs two distinct market-share thresholds, both *sufficient* conditions that are conservative. The *supermodularity threshold* $\bar{s}^* \equiv (\alpha - 1)/(2\alpha)$ from Proposition 6 governs per-market convexity under the CES benchmark and depends only on $(\sigma, \theta)$. The *pricing-contraction threshold* $\bar{s}(\sigma, \eta, F)$ from Corollary 17

Table 1: Market-share threshold $\bar{s}(\sigma, \eta, F)$ for BNE uniqueness

| | $\kappa < 1$ ($F = 2$) | | | | $\kappa < 1/7$ ($F = 8$) | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma \backslash \eta$ | 1.5 | 2.0 | 3.0 | 4.0 | 1.5 | 2.0 | 3.0 | 4.0 |
| 3 | 0.843 | 1.000 | — | — | 0.386 | 0.532 | — | — |
| 5 | 0.687 | 0.795 | 1.000 | 1.000 | 0.283 | 0.339 | 0.482 | 0.756 |
| 8 | 0.616 | 0.675 | 0.802 | 0.953 | 0.232 | 0.265 | 0.333 | 0.411 |
| 10 | 0.593 | 0.640 | 0.736 | 0.841 | 0.214 | 0.241 | 0.294 | 0.350 |

governs pricing uniqueness via $\|\Lambda\|_\infty$ and additionally depends on $(\eta, F)$. At the baseline $(\sigma, \eta, \theta) = (5, 2, 3)$ with $F = 8$: $\bar{s}^* = 0.125$ and $\bar{s} \approx 0.339$, while $\max_{n,f} s_{nf} = 0.39$—both thresholds are violated. Nevertheless, aggregate supermodularity holds (Jia–Tarski gap = 0; Remark 7) and pricing uniqueness holds under the tighter spectral condition ($\rho(\Lambda) = 0.096 \ll 1$). The share thresholds are useful as closed-form diagnostics but are not necessary conditions.

**TFP dispersion and marginal sites.** Fix all parameters except TFP and define the dispersion ratio $\delta = a_{\max}/a_{\min}$. As $\delta$ increases, the number of marginal sites weakly decreases: in the limit $\delta \to \infty$, only the highest-TFP firms are active in any market, and $\kappa \to 0$ for non-dominant firms. When $a_f = a$ for all $f$, market shares are balanced at approximately $s \approx 1/F$. Since $\kappa_{nf}$ is increasing and convex in $s_{nf}$, balanced shares actually *minimize* the feedback coefficient: condition (3.3) is *easier* to satisfy under identical TFP. Multiplicity in this case requires marginal sites driven by the available-set structure $\{\mathcal{A}_f\}$, which is absent at the baseline. Section 5.5, Experiment 3 confirms this prediction numerically.

## 3.3 Example: Dimensionality Reduction through Combinatorial Structure

I illustrate how the CDCP approach reduces dimensionality relative to the binary-entry framework of Espin-Sanchez et al. (2023) using a small economy (hereafter the *small calibration*): $F = 4$ firms, $N = 3$ markets, with two "strong" firms ($a_f = 1.5, 1.2$; access to all 3 sites) and two "weak" firms ($a_f = 1.0, 0.8$; access to 2 sites each). The complete-information equilibrium is unique (supremum and infimum Jia–Tarski iterates coincide) with all firms active at their full site sets.

**The Espin-Sanchez et al. approach.** Treating each firm-site pair $(i, f)$ as an independent binary player gives $M = 10$ players in a threshold system $\nu_{if} = G_0(\bar{\varepsilon}_{if}/\sigma_\varepsilon)$. The configuration space has $2^{10} = 1{,}024$ joint action profiles. Each weak firm has one clearly profitable site (marginal value 0.28–0.71 from empty) and one marginal or unprofitable site (MV $\approx -0.003$ to $-0.14$ from empty). Under Espin-Sanchez et al., these marginal individual-site decisions are independent threshold variables that expand the fixed-point dimension.

**The CDCP approach.** Recognizing that each firm solves a combinatorial problem over $2^{|\mathcal{A}_f|}$ configurations collapses the 10-player game into a 4-player game. Within-firm complementarity bundles the weak firms' marginal sites with their profitable sites: for both weak firms, the Fréchet cost complementarity creates positive increasing differences (ID = 0.020 and 0.044), so that each site is profitable conditional on the other even when individually marginal. The CDCP optimal set is the full site set $\mathcal{A}_f$ for each firm, and the 4-player threshold system has 60% fewer degrees of freedom.

**Takeaway.** The combinatorial structure collapses $M = \sum_f |\mathcal{A}_f|$ binary decisions into $F$ firm-level decisions—a 60% reduction in threshold dimension for this example. This reduction is not merely computational: it bundles individually marginal site decisions into clearly-determined firm-level packages via within-firm cost complementarity, so that each firm's activation is well-determined even when individual sites are near the profitability threshold. At this parameterization, the Espin-Sanchez et al. system is also unique, so the benefit here is computational rather than equilibrium-theoretic. However, the lower-dimensional threshold system yields a tighter Jacobian bound in the contraction analysis of Section 3.2, because within-firm complementarities are internalized rather than treated as additional cross-player interactions. Numerical details are in the companion script `main_example_2x2.m`.

## 4. Computational Algorithm

The equilibrium defined in Section 2 is computed by a three-layer nested algorithm. I describe each layer and present pseudocode; Appendix E reports performance diagnostics

for four algorithm variants and a 40-country extension demonstrating scalability. Figure 2 illustrates the structure.

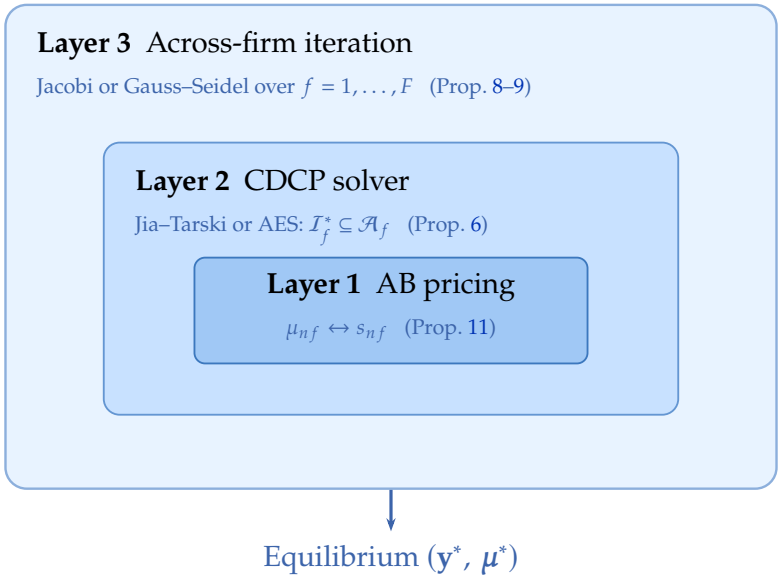


Figure 2: Three-layer algorithm nesting. Each profit evaluation at Layer 2 calls the AB pricing fixed point (Layer 1); the across-firm iteration at Layer 3 updates rival demand indices until convergence. Four algorithm variants arise from the $2 \times 2$ choice of solver at Layers 2 and 3 (Appendix E).

## 4.1 Three-Layer Structure

**Layer 1 (Innermost): Atkeson–Burstein markup fixed point.** For given marginal cost $c_{nf}$ and rival demand index $D_{-f}(n)$, the equilibrium markup $\mu_{nf}$ is the fixed point of:

$$\mu_{nf}^{(t+1)} = \frac{\epsilon_{nf}^{(t)}}{\epsilon_{nf}^{(t)} - 1}, \qquad \frac{1}{\epsilon_{nf}^{(t)}} = \frac{1 - s_{nf}^{(t)}}{\sigma} + \frac{s_{nf}^{(t)}}{\eta}, \tag{4.1}$$

where $s_{nf}^{(t)} = (\mu_{nf}^{(t)} c_{nf})^{1-\sigma} / \left[(\mu_{nf}^{(t)} c_{nf})^{1-\sigma} + D_{-f}(n)\right]$. The iteration is initialized at the CES markup $\mu^{(0)} = \sigma/(\sigma - 1)$ and damped with weight $\omega_\mu = 0.5$:

$$\mu^{(t+1)} \leftarrow (1 - \omega_\mu)\, \mu^{(t)} + \omega_\mu\, \mu^{(t+1)}.$$

Convergence is guaranteed by Proposition 11 whenever $\rho(\Lambda) < 1$.

**Layer 2 (Middle): Single-agent CDCP.** Given the rival demand index $D_{-f}(n)$ for all markets $n$, each firm $f$ solves the combinatorial optimization:

$$\mathcal{I}_f^* \in \arg\max_{\mathcal{I}_f \subseteq \mathcal{A}_f} \pi_f\left(\mathcal{I}_f; D_{-f}\right), \tag{4.2}$$

where each profit evaluation calls the AB markup fixed point (Layer 1). Two inner solvers can be used:

(a) **Jia–Tarski bounds** (Jia, 2008): Compute an *upper bound* $\overline{\mathcal{I}}_f$ by starting from all candidates $\mathcal{A}_f$ and iteratively removing elements with negative marginal value; compute a *lower bound* $\underline{\mathcal{I}}_f$ by starting from $\emptyset$ and iteratively adding elements with positive marginal value. If $\overline{\mathcal{I}}_f = \underline{\mathcal{I}}_f$, the solution is exact.

(b) **AES branch-and-bound** (Arkolakis et al., 2023): Alternate between a *squeezing phase* (simultaneously squeeze included/excluded sets by evaluating marginal values against both the lower and upper bounds of the candidate set) and a *branching phase* (recursively split on an undecided candidate). The squeezing phase typically resolves most candidates, leaving few branches.

Both solvers find the exact optimum. When the equilibrium shares satisfy the conditions of Proposition 6, the Jia–Tarski bounds converge without residual gap, computationally certifying that supermodularity holds at the equilibrium.

**Layer 3 (Outermost): Cross-firm iteration.** The rival demand index $D_{-f}(n) = \sum_{f' \neq f} (\mu_{nf'} c_{nf'})^{1-\sigma}$ depends on all firms' equilibrium prices, which in turn depend on their activation sets. Two outer iteration methods resolve this fixed point:

(a) **Jacobi** (simultaneous best response): In each round $k$, every firm $f$ computes its best-response activation set $\mathcal{I}_f^{(k+1)}$ against the rival demand $D_{-f}^{(k)}$ from the *previous* round. All firms update simultaneously, and price-power contributions are damped:

$$p_{nf}^{(k+1)} \leftarrow (1-\omega)\, p_{nf}^{(k)} + \omega\, \tilde{p}_{nf}^{(k+1)},$$

where $\tilde{p}_{nf}^{(k+1)}$ is the price implied by $\mathcal{I}_f^{(k+1)}$ and $\omega \in (0, 1]$ is a damping parameter. Damping does not affect the set of fixed points—only the convergence path—because

any fixed point satisfies $\tilde{p} = p$.

(b) **Gauss–Seidel** (sequential best response): Firms are processed in a fixed order $(f_1, \ldots, f_F)$. When firm $f_j$ updates, it uses the most recent price-power contributions from firms $f_1, \ldots, f_{j-1}$ (already updated in this round) and the previous round's contributions from $f_{j+1}, \ldots, f_F$.

By Proposition 9, Jacobi iteration converges monotonically to extremal equilibria when initialized from the supremum or infimum; Gauss–Seidel convergence is empirically robust (Remark 10).

**Algorithmic recommendation.** Propositions 6–9 yield a concrete recipe: initialize the outer iteration from the *supremum* (all available sites active) and solve each firm's CDCP via the Jia–Tarski or AES algorithm. Starting from the supremum guarantees convergence to the upper bound of the equilibrium set (Proposition 9), which at all tested parameterizations is itself the unique Nash equilibrium. The nested CES structure further accelerates computation: the interaction matrix $\Lambda$ is block-diagonal across markets ($N$ blocks of size $F \times F$), so $\rho(\Lambda) = \max_n \rho(\Lambda_n)$ is computable in $O(NF^3)$, and the closed-form feedback coefficient (3.2) reduces the contraction check to a single market-share comparison.

## 4.2 Pseudocode

Algorithm 1 summarizes the full procedure.

# 5. Numerical Illustration

I illustrate the theoretical results with a baseline economy featuring $N = 20$ countries and $F = 8$ firms. All four solver variants from Section 4 yield the same equilibrium; performance diagnostics are in Appendix E. The analysis is organized around three questions: does endogenous pricing matter quantitatively (Section 5.2)? Is the BNE contraction modulus below one (Section 5.3)? And how robust is uniqueness across the parameter space (Section 5.5)?

**Algorithm 1** Three-Layer Equilibrium Solver

**Input:** Parameters $(\sigma, \eta, \theta, \{a_f\}, \{\phi_i, w_i, \tau_{in}\}, \{\Omega_n\})$; inner solver $\in$ {JT, AES}; outer method $\in$ {Jacobi, GS}
**Output:** Equilibrium $(\{\mathcal{I}_f^*\}, \{\mu_{nf}^*\}, \{s_{nf}^*\})$

**Initialize:** $\mathcal{I}_f^{(0)} \leftarrow \mathcal{A}_f$ for all $f$; compute $p_{nf}^{(0)}$ at CES markups
**for** $k = 0, 1, 2, \ldots$ **do**
  **for** $f = 1, \ldots, F$ **do** ▷ Jacobi: use $D^{(k)}$; GS: use latest $D$
    $D_{-f}(n) \leftarrow \sum_{f' \neq f} [p_{nf'}^{1-\sigma}]^{\text{current}}$ for all $n$
    $\mathcal{I}_f^{(k+1)} \leftarrow \textsc{InnerSolver}(f, D_{-f}, \text{par})$ ▷ Layer 2: CDCP
    Update $p_{nf}$ via $\textsc{ABMarkup}(c_{nf}(\mathcal{I}_f^{(k+1)}), D_{-f})$ ▷ Layer 1
  **end for**
  **if** $\max_{n,f} |p_{nf}^{(k+1)} - p_{nf}^{(k)}| <$ `tol` **then break**
  **end if**
**end for**

## 5.1 Parameterization

Table 2 summarizes the common parameters, chosen to satisfy the supermodularity condition $(\sigma - 1)/\theta = 4/3 > 1$. The within-sector elasticity $\sigma = 5$ lies in the middle of the range estimated by Broda and Weinstein (2006) for manufacturing sectors. The across-sector elasticity $\eta = 2$ is standard in multi-sector trade models. The Fréchet shape $\theta = 3$ is conservative: Eaton and Kortum (2002) estimate $\theta \approx 8.28$, and Simonovska and Waugh (2014) report $\theta \in [4, 5]$; the lower value makes the supermodularity condition $(\sigma-1)/\theta > 1$ the most binding, providing a stringent test of the theoretical results.

Table 2: Common Parameters

| Parameter | Description | Value |
|---|---|---|
| $\sigma$ | Within-sector elasticity | 5 |
| $\eta$ | Across-sector elasticity | 2 |
| $\theta$ | Fréchet shape | 3 |
| $(\sigma - 1)/\theta$ | Supermodularity ratio | 4/3 |
| $\phi_i$ | Fixed cost per site | 0.08–0.15 |

**Sensitivity to $\theta$.** Referees may note that $\theta = 3$ is below standard empirical estimates. Table 3 reports equilibrium diagnostics at $\theta \in \{3, 4, 5, 8\}$, spanning the range from our baseline to the Eaton and Kortum (2002) estimate. At $\theta \geq 4$, the ratio $(\sigma - 1)/\theta \leq 1$ and

the analytic supermodularity condition in Proposition 6(i) fails entirely. Nevertheless, the Jia–Tarski gap equals zero at every $\theta$, confirming that the computational certification in Proposition 6(ii) succeeds. The spectral radius $\rho(\Lambda) < 0.1$ throughout, so pricing uniqueness is not a binding constraint. Active sites decline with $\theta$ (from 37 to 15) because lower Fréchet dispersion reduces the profitability of marginal locations, but uniqueness is preserved.

Table 3: Uniqueness Diagnostics across $\theta$

| $\theta$ | $(\sigma-1)/\theta$ | $\bar{s}^*$ | max $s$ | JT gap | $\rho(\Lambda)$ | Active sites |
|---|---|---|---|---|---|---|
| 3 | 1.333 | 0.125 | 0.394 | 0 | 0.096 | 37 |
| 4 | 1.000 | — | 0.393 | 0 | 0.096 | 24 |
| 5 | 0.800 | — | 0.393 | 0 | 0.095 | 16 |
| 8 | 0.500 | — | 0.393 | 0 | 0.095 | 15 |

*Notes:* Baseline economy ($N = 20$, $F = 8$) solved from both supremum and infimum starting points. $\bar{s}^*$: share threshold for per-market supermodularity; "—" indicates the analytic condition fails ($(\sigma-1)/\theta \leq 1$). JT gap: maximum Jia–Tarski gap across all firms. $\rho(\Lambda)$: spectral radius of the pricing interaction matrix. All configurations yield a unique equilibrium.

**Broader parameter sensitivity.** Table 4 extends the analysis to a $3 \times 3 \times 4$ grid over $(\sigma, \eta, \theta)$, spanning $\sigma \in \{3, 5, 8\}$, $\eta \in \{1.5, 2, 3\}$, and $\theta \in \{3, 4, 5, 8\}$. The ratio $(\sigma-1)/\theta$ ranges from 0.25 to 2.33; in 18 of 28 configurations the analytic supermodularity condition fails ($(\sigma-1)/\theta \leq 1$). Nevertheless, the Jia–Tarski gap equals zero and the equilibrium is unique at *all* 28 configurations. The spectral radius $\rho(\Lambda)$ increases with $\sigma$ and decreases with $\eta$—reaching 0.225 at $(\sigma, \eta) = (8, 1.5)$—but remains well below one throughout. These results confirm that the uniqueness findings are not an artifact of the baseline parameterization.

The 20 countries are organized in four clusters of five: *advanced* ($n = 1$–$5$, $\bar{w} \approx 1.1$, $\bar{R} \approx 12$), *upper-middle* ($n = 6$–$10$, $\bar{w} \approx 0.6$, $\bar{R} \approx 7$), *lower-middle* ($n = 11$–$15$, $\bar{w} \approx 0.3$, $\bar{R} \approx 4.5$), and *frontier* ($n = 16$–$20$, $\bar{w} \approx 0.15$, $\bar{R} \approx 2.5$). Wages and expenditures are drawn with within-cluster variation. Trade costs follow a three-tier block structure: $\tau = 1.15$ within cluster, $\tau = 1.3$ between adjacent clusters, $\tau = 1.5$ between distant clusters. Fixed costs vary by cluster: $\phi_i = 0.15$ (advanced), 0.12 (upper-middle), 0.10 (lower-middle), 0.08 (frontier), reflecting higher operating costs in developed markets.

The eight firms are designed to create a hierarchy from global to regional players,

Table 4: Uniqueness Diagnostics across $(\sigma, \eta, \theta)$

| | | $\eta = 1.5$ | | | | $\eta = 2$ | | | | $\eta = 3$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | $\theta$ | $\frac{\sigma-1}{\theta}$ | $\rho$ | JT | sites | $\frac{\sigma-1}{\theta}$ | $\rho$ | JT | sites | $\frac{\sigma-1}{\theta}$ | $\rho$ | JT | sites |
| 3 | 3 | 0.67 | .047 | 0 | 39 | 0.67 | .025 | 0 | 38 | | | | |
| 3 | 5 | 0.40 | .047 | 0 | 19 | 0.40 | .026 | 0 | 18 | | | | |
| 3 | 8 | 0.25 | .048 | 0 | 16 | 0.25 | .026 | 0 | 16 | | | | |
| 5 | 3 | 1.33 | .125 | 0 | 38 | 1.33 | .096 | 0 | 37 | 1.33 | .053 | 0 | 36 |
| 5 | 5 | 0.80 | .123 | 0 | 17 | 0.80 | .095 | 0 | 16 | 0.80 | .052 | 0 | 16 |
| 5 | 8 | 0.50 | .123 | 0 | 15 | 0.50 | .095 | 0 | 15 | 0.50 | .052 | 0 | 15 |
| 8 | 3 | 2.33 | .225 | 0 | 39 | 2.33 | .208 | 0 | 36 | 2.33 | .160 | 0 | 34 |
| 8 | 5 | 1.40 | .222 | 0 | 18 | 1.40 | .204 | 0 | 18 | 1.40 | .157 | 0 | 18 |
| 8 | 8 | 0.88 | .222 | 0 | 15 | 0.88 | .204 | 0 | 14 | 0.88 | .157 | 0 | 13 |

*Notes:* Baseline economy ($N = 20$, $F = 8$). $\sigma > \eta$ required; $(\sigma, \eta) = (3, 3)$ excluded. Rows with $\theta = 4$ omitted for compactness (available on request). JT: Jia–Tarski gap (0 at all configurations). $\rho$: spectral radius $\rho(\Lambda)$. All 28 configurations yield a unique equilibrium from both supremum and infimum starting points.

generating the market-share heterogeneity needed to stress-test the pricing contraction condition:

| Firm | TFP $a_f$ | Available Sites $\mathcal{A}_f$ | $2^{\lvert\mathcal{A}_f\rvert}$ |
|---|---|---|---|
| 1 | 2.0 | $\{1, \ldots, 20\}$ | 1,048,576 |
| 2 | 1.8 | $\{1, \ldots, 20\}$ | 1,048,576 |
| 3 | 1.5 | $\{1, \ldots, 15\}$ | 32,768 |
| 4 | 1.5 | $\{6, \ldots, 20\}$ | 32,768 |
| 5 | 1.2 | $\{1, \ldots, 10\}$ | 1,024 |
| 6 | 1.2 | $\{11, \ldots, 20\}$ | 1,024 |
| 7 | 1.0 | $\{1, \ldots, 5\}$ | 32 |
| 8 | 1.0 | $\{16, \ldots, 20\}$ | 32 |

More productive firms (higher $a_f$) tend to have wider geographic reach. Firms 1 and 2 are global players with access to all 20 sites, while Firms 7 and 8 are regional players restricted to a single cluster. Exhaustive enumeration of $2^{20} \approx 10^6$ subsets is computationally expensive; the Jia–Tarski and AES algorithms exploit the lattice structure to solve each firm's CDCP in polynomial time. For the contraction-modulus computation in Section 5.3, I also use a small economy ($N = 3$, $F = 4$, $M = 12$) where exact enumeration of all rival configurations is feasible.

## 5.2 Variable versus Constant Markups

To quantify the role of endogenous markups, I solve the equilibrium twice—once with Atkeson–Burstein (AB) variable markups and once with constant CES markups $\mu = \sigma/(\sigma - 1) = 1.25$—holding all other parameters fixed at baseline values.

**Markups and market shares.** Figure 3(a) plots each (firm, market) pair's AB markup against its market share, showing the monotone relationship implied by the Atkeson–Burstein pricing rule: firms with higher shares face lower perceived demand elasticities, inflating their markups above the CES floor of $\sigma/(\sigma - 1) = 1.25$. The cross-sectional mean is $\bar{\mu}^{AB} = 1.36$ with a range of $[1.27, 1.47]$, versus the uniform $\mu^{CES} = 1.25$. This markup inflation compresses market shares: the maximum share under AB is 0.39, compared with 0.47 under CES—a 17% reduction (Figure 3b, where all points lie below the 45-degree line). The compression is the intensive-margin channel through which endogenous markups generate pro-competitive effects.

**Activation.** The markup difference propagates to the extensive margin: under AB markups, five active firms operate 37 sites in total; under CES markups, only 32 (Figure 3c). The five additional sites are marginal locations where variable markup revenue covers fixed costs, but constant-markup revenue does not. The activation difference is concentrated in mid- and low-TFP firms, each losing one to two sites under CES.

**Profits.** Despite higher markups per unit, the markup-share compression does not dominate: AB aggregate profits exceed CES profits by 30.2% (Figure 3d). The profit gap widens as firm TFP falls (from −26% to −39% when switching from AB to CES), because lower-TFP firms lose proportionally more activation revenue. The profit difference is economically meaningful and underscores the bias from imposing constant markups in quantitative trade and multinational production models. From a consumer welfare perspective, the two pricing rules have offsetting effects: the 16% activation expansion under AB generates additional product variety (lowering $P_n$), but the higher variable markups ($\bar{\mu}^{AB} = 1.36$ vs. $\mu^{CES} = 1.25$) raise prices. Computing the CES real-income index $W_n = R_n/P_n$ at each market, AB welfare is on average 11% *lower* than CES welfare. Decomposing: the *variety channel* (37 vs. 32 active sites under AB vs. CES) lowers the price index and raises

welfare, whereas the *markup channel* ($\bar{\mu}^{AB} = 1.36$ vs. $\mu^{CES} = 1.25$) raises prices and reduces welfare. At all 20 markets, the markup channel exceeds the variety channel, yielding a net welfare loss that is largest in advanced markets (where concentration is highest) and smallest in frontier markets.[2] The welfare decomposition underscores that the activation decision—the dimension on which this article's uniqueness result operates—has quantitatively significant welfare consequences that constant-markup models misattribute to other channels.

**Pricing interaction.** The spectral radius of the markup interaction matrix is $\rho(\Lambda) = 0.096$ under AB, confirming that pricing feedback is quantitatively small even though it is qualitatively important for the activation margin. Note that the maximum market share of 0.39 exceeds the infinity-norm threshold $\bar{s} = 0.339$ from Corollary 17, so the sufficient condition $\max \kappa < 1/(F-1)$ is violated; pricing uniqueness instead follows from the tighter spectral condition $\rho(\Lambda) < 1$. Under CES, $\rho(\Lambda) = 0$ by construction—there is no markup feedback. The spectral condition resolves the intensive margin (pricing); whether the *full game*—including the extensive margin—also has a unique equilibrium is the subject of Section 5.3.

## 5.3 Contraction Modulus: The Bathtub

Proposition 13(iv) predicts that the contraction modulus $\gamma(\sigma_\varepsilon) = \sup_\nu \|J_T(\nu)\|_\infty$ has a bathtub shape, vanishing as $\sigma_\varepsilon \to 0$ and $\sigma_\varepsilon \to \infty$. I verify this prediction numerically.

For each $\sigma_\varepsilon/\bar{\phi}$ on a log-spaced grid (where $\bar{\phi}$ is the mean fixed cost), I solve the BNE fixed point $\nu^* = T(\nu^*)$ iteratively and compute $\gamma$ via the Jacobian factorization (3.7). I report results for the small economy ($N = 3$, $F = 4$) with exact enumeration of rival configurations over a 40-point grid spanning $\sigma_\varepsilon/\bar{\phi} \in [0.003, 16]$.

Figure 4 confirms the bathtub shape. At both extremes, $\gamma$ vanishes as predicted by Proposition 13(iv): exponentially fast as $\sigma_\varepsilon \to 0$ (the standardized thresholds diverge) and at rate $1/\sigma_\varepsilon$ as $\sigma_\varepsilon \to \infty$. The maximum contraction modulus is $\gamma_{\max} = 0.951$ at $\sigma_\varepsilon/\bar{\phi} \approx 0.60$, well below one, confirming that the BNE is unique for all noise scales tested. The peak occurs at an intermediate $\sigma_\varepsilon$ where the density dampening $g_0(z)/\sigma_\varepsilon$ is maximized, consistent with the analytic decomposition $\gamma \leq (\|g_0\|_\infty/\sigma_\varepsilon) \cdot \|J_v\|_\infty$.

[2] The welfare comparison is computed holding $R_n$ fixed at its initial value (Remark 1).

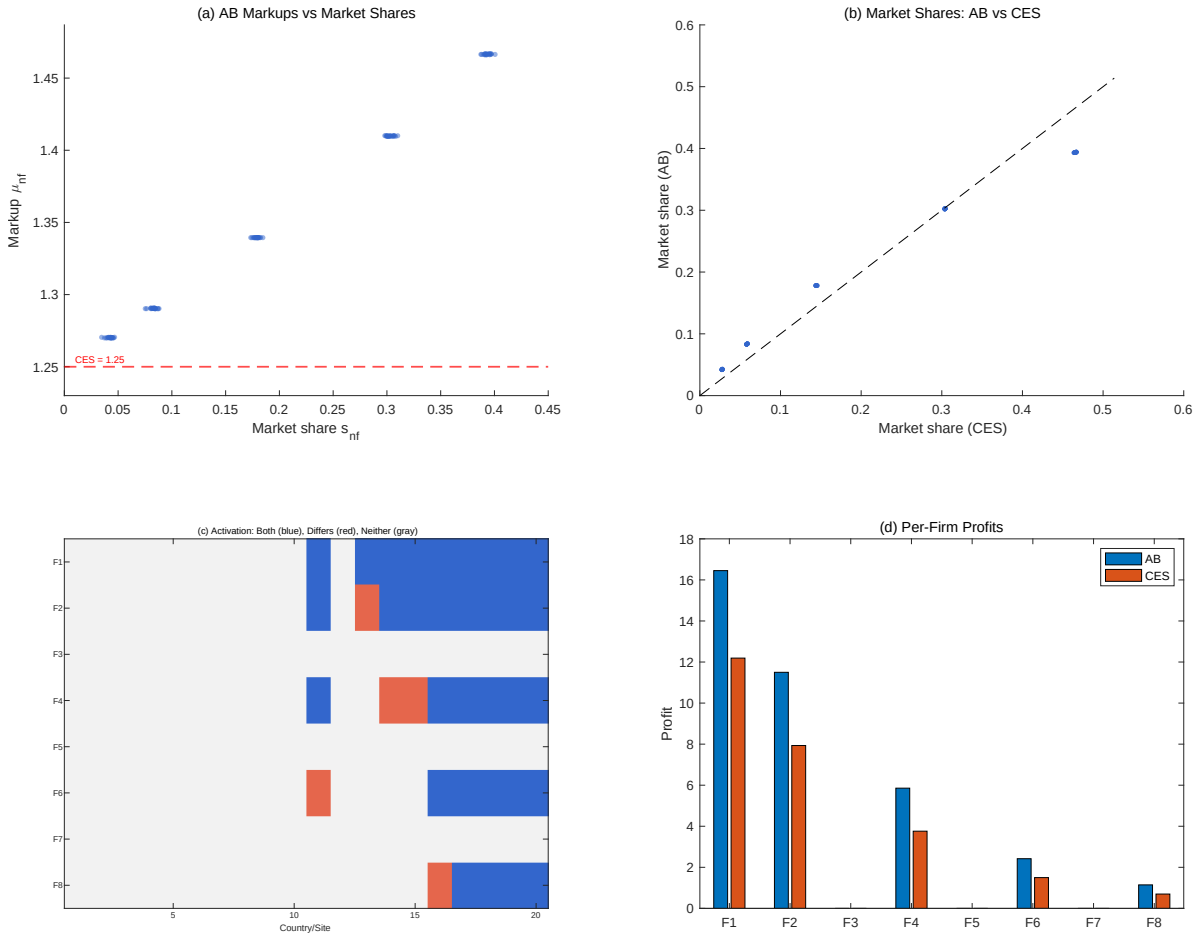


Figure 3: AB variable markups versus constant CES markups. (a) AB markups versus market shares: each dot is a (firm, market) pair; the dashed line marks the CES constant markup $\sigma/(\sigma-1) = 1.25$. (b) Market shares under both models; the 45-degree line indicates equality. (c) Activation heatmap: blue = active under both; red = active under one model only; gray = inactive. (d) Per-firm profits.

**Baseline verification.** For the baseline economy ($N = 20, F = 8, M = 100$ firm–site pairs), exact computation of $\|J_T\|_\infty$ is intractable: Monte Carlo estimation of the Jacobian suffers from noise amplification by the factor $g_0(z_k)/(w_k\sigma_\varepsilon)$, which reaches $O(10^2)$ at threshold entries, causing the accumulated error across $M = 100$ columns to overwhelm the signal.

I instead verify BNE uniqueness directly by solving the fixed-point iteration from three diverse initializations—the complete-information supremum, the infimum, and a uniform starting point ($\nu = 0.5$)—at each $\sigma_\varepsilon$ on a 15-point grid spanning $\sigma_\varepsilon/\bar{\phi} \in [0.01, 10]$. All three initializations converge to the same BNE at every grid point (maximum $\ell_\infty$ distance $< 0.004$, consistent with the Monte Carlo noise floor), providing strong empirical evidence of uniqueness throughout the baseline parameterization.

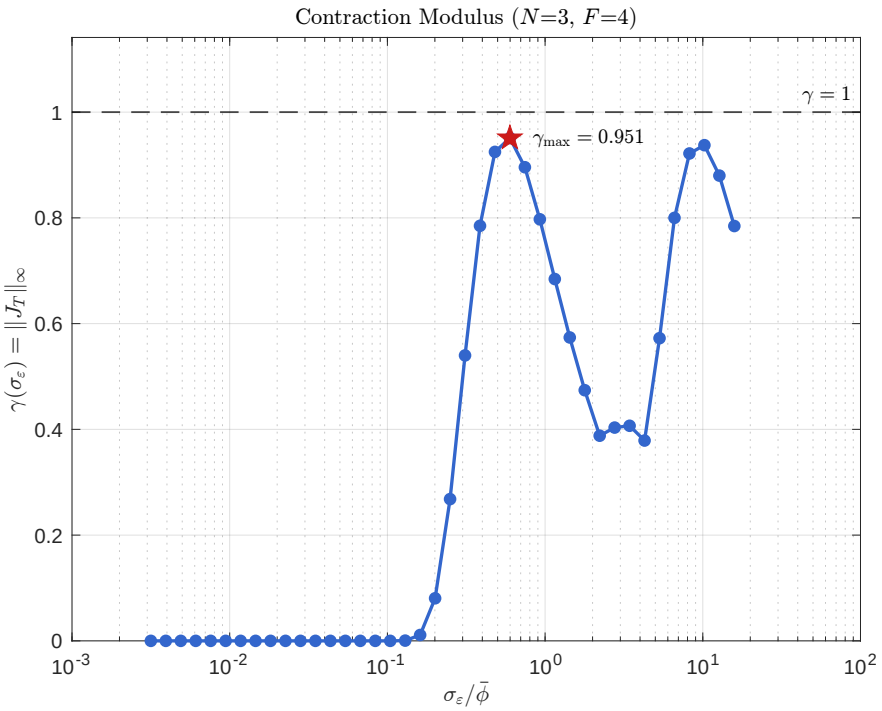


Figure 4: Contraction modulus $\gamma(\sigma_\varepsilon)$ for the small economy ($N = 3$, $F = 4$). The bathtub shape confirms Proposition 13(iv): $\gamma \to 0$ as $\sigma_\varepsilon \to 0$ (exponential density decay) and as $\sigma_\varepsilon \to \infty$ ($1/\sigma_\varepsilon$ dampening). The peak $\gamma_{\max} = 0.951$ (marked) is well below the dashed $\gamma = 1$ threshold.

**Scaling with $F$ and $N$.** A natural concern is whether the peak contraction modulus $\gamma_{\max}$ grows with the number of firms $F$ or sites $N$, potentially exceeding one in larger economies. To investigate, I compute the exact bathtub curve for a grid of small economies with $F \in \{2, 3, 4, 5, 6\}$ and $N \in \{2, 3, 4\}$, subject to $M = \sum_f |\mathcal{A}_f| \leq 12$ (required for exact Jacobian computation via finite differences). For each $(F, N)$, I construct a symmetric economy with dispersed TFP ($a_f \in [0.8, 1.5]$), dispersed wages ($w_i \in [0.3, 1.0]$), and iceberg trade costs ($\tau = 1.3$ off-diagonal).

Figure 5 presents the results. The bathtub shape is preserved across all configurations. Two opposing forces determine $\gamma_{\max}$: (i) $\gamma_{\max}$ *decreases* monotonically with $F$ for fixed $N$—more firms means smaller market shares and weaker marginal-value sensitivity—falling from 3.65 at $(F=2, N=2)$ to 0.35 at $(F=6, N=2)$; (ii) $\gamma_{\max}$ *increases* with $N$ for fixed $F$—more sites create additional near-marginal activation decisions. For the $N = 2$ cross-section, $\gamma_{\max} < 1$ at $F \geq 4$, whereas for $N \geq 3$ the sufficient condition may be violated at small $F$. The contraction condition is thus most easily satisfied in the empirically relevant regime of many firms with moderate market shares—precisely the setting of the baseline economy.

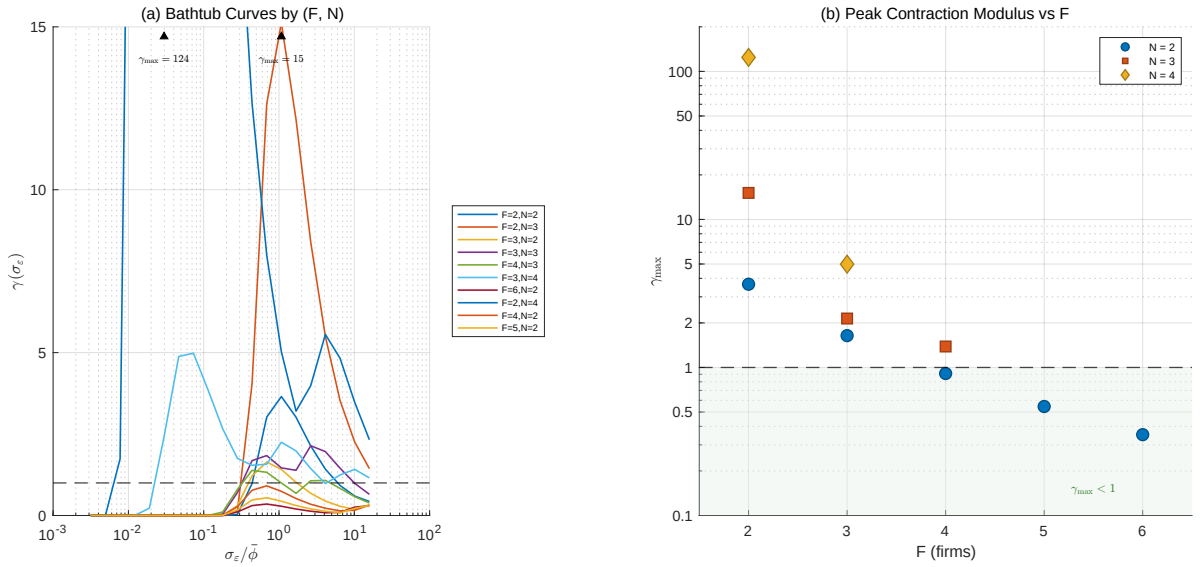


Figure 5: Scaling of the contraction modulus with $(F, N)$. (a) Bathtub curves $\gamma(\sigma_\varepsilon)$ for economies of different sizes. (b) Peak contraction modulus $\gamma_{\max}$ vs. number of firms $F$, with color indicating $N$.

Whether $\gamma_{\max} < 1$ at a given parameterization depends on the full vector of primitives, not only on $(F, N)$; the bathtub decomposition (Proposition 13) provides a diagnostic that can be checked for any specific calibration.

**BNE at a multiplicity point.** With TFP dispersion ratio $a_{\max}/a_{\min} = 3$ in the small economy, the complete-information game admits two Nash equilibria differing by one marginal site. Figure 6(a) traces the activation probabilities: the marginal site transitions smoothly from $\nu \approx 0$ at small noise to $\nu \approx 0.49$ at large noise, while non-marginal sites remain near $\{0, 1\}$. Figure 6(b) shows $\gamma_{\max} \approx 5.5 > 1$: the sufficient condition of Theorem 16(e) is violated, yet the BNE iteration converges from all initializations, confirming that the infinity-norm bound is conservative at the multiplicity boundary. As $\sigma_\varepsilon \to 0$, the BNE selects the equilibrium obtained by initializing from the supremum—an empirical regularity consistent with the lattice structure of Proposition 8 but not formally proved in general.

## 5.4 Equilibrium Uniqueness at Baseline

Figure 7 presents the *site activation map*, classifying each (firm, site) pair as active (blue), inactive but available (white), or unavailable (gray). All four solver combinations—two

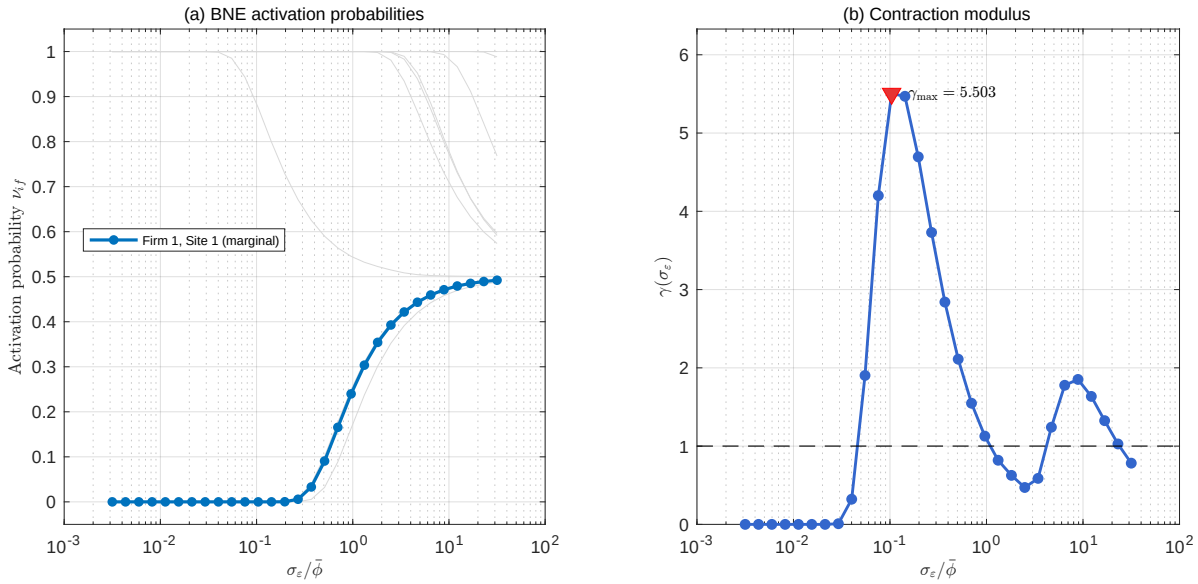


Figure 6: BNE at a complete-information multiplicity point ($N = 3$, $F = 4$, $a_{\max}/a_{\min} = 3$). (a) Activation probabilities: the marginal site (firm 1, site 1; blue) transitions from $\nu \approx 0$ to $\nu \approx 0.49$ as $\sigma_\varepsilon$ increases; non-marginal sites (gray) remain near $\{0, 1\}$. (b) Contraction modulus: the bathtub shape is preserved but $\gamma_{\max} > 1$, so the sufficient condition is not met at intermediate noise.

inner solvers (Jia–Tarski and AES) from two starting points (supremum and infimum)—converge to the same activation profile, confirming uniqueness at the baseline parameterization. Five of eight firms are active, activating 37 sites in total with aggregate profit $\sum \pi_f = 37.37$. Firms 3, 5, and 7 activate no sites—they are crowded out by higher-TFP competitors with wider production networks. Active firms concentrate in the lower-middle and frontier clusters where labor is cheapest, leaving the advanced-economy clusters inactive because of high operating costs. The absence of any marginal sites (sites that switch between equilibria) confirms uniqueness: the upper and lower bounds of Proposition 8 coincide, and both are Nash equilibria.

**Interaction diagnostics.** The pricing sub-problem is strongly contractive: $\rho(\Lambda) = 0.096 \ll 1$ (Proposition 11), so the intensive margin is resolved. The extensive margin is also resolved at baseline: every site's activation profit gap $\Delta\pi_{if}$ is bounded away from zero, so no site is at the margin of profitability.

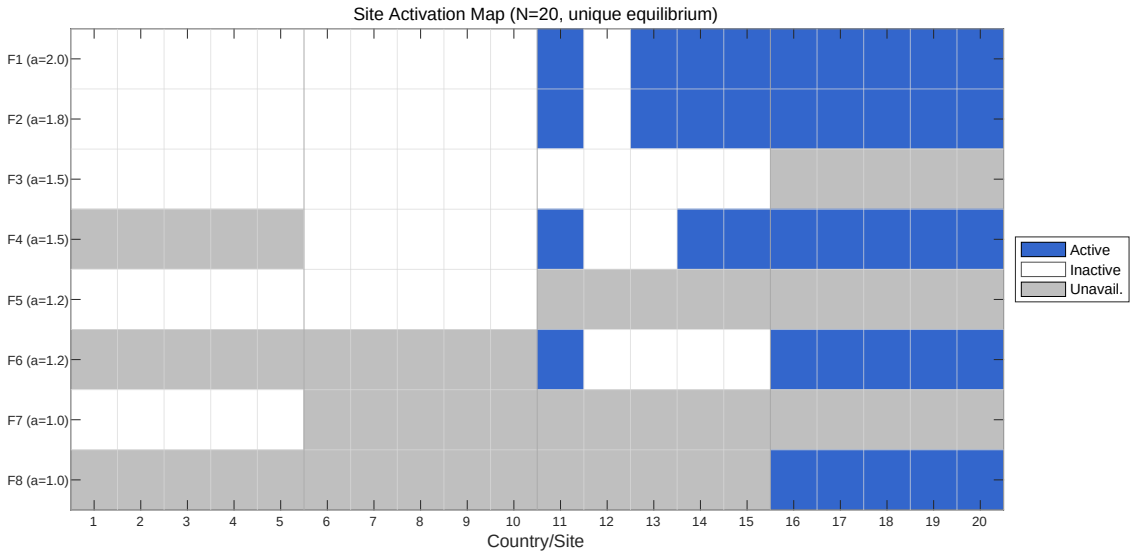


Figure 7: Site Activation Map ($N = 20$, $F = 8$, unique equilibrium). Blue: active sites; white: available but inactive; gray: unavailable.

## 5.5 Robustness of Uniqueness

I evaluate robustness along three dimensions: TFP heterogeneity, fixed-cost levels, and pricing feedback intensity. For each experiment, I compute the feedback coefficient $\kappa_{nf}$, the spectral radius $\rho(\Lambda)$, and run all four solver combinations from both starting points to detect multiplicity.

Table 5: Uniqueness Diagnostics ($N = 20$, $F = 8$)

| Experiment | $\max \kappa$ | $\frac{1}{F-1}$ | $\rho(\Lambda)$ | $n_{\text{eq}}$ | Marginal sites |
|---|---|---|---|---|---|
| Baseline (het. TFP, het. $\phi$) | 0.193† | 0.143 | 0.096 | 1 | 0 |
| Exp. 1: Identical TFP | 0.052 | 0.143 | 0.052 | 1 | 0 |
| Exp. 2: Uniform $\phi$ | 0.193 | 0.143 | 0.097 | 1 | 0 |

† At baseline, $\max \kappa = 0.193 > 1/(F-1) = 0.143$: the infinity-norm condition of Proposition 11 is violated due to the dominant firms' high shares. However, the tighter spectral condition $\rho(\Lambda) = 0.096 < 1$ confirms pricing uniqueness.

Table 5 reports the results. At baseline, $\rho(\Lambda) = 0.096 \ll 1$, confirming that the pricing sub-problem is strongly contractive (Proposition 11). The equilibrium is unique with zero marginal sites, so the extensive margin is also resolved.

**Experiments 1–2: Symmetric firms and uniform costs.** Setting $a_f = 1.40$ for all $f$ (Experiment 1) reduces $\max \kappa$ from 0.193 to 0.052 and $\rho(\Lambda)$ from 0.096 to 0.052 (Table 5),

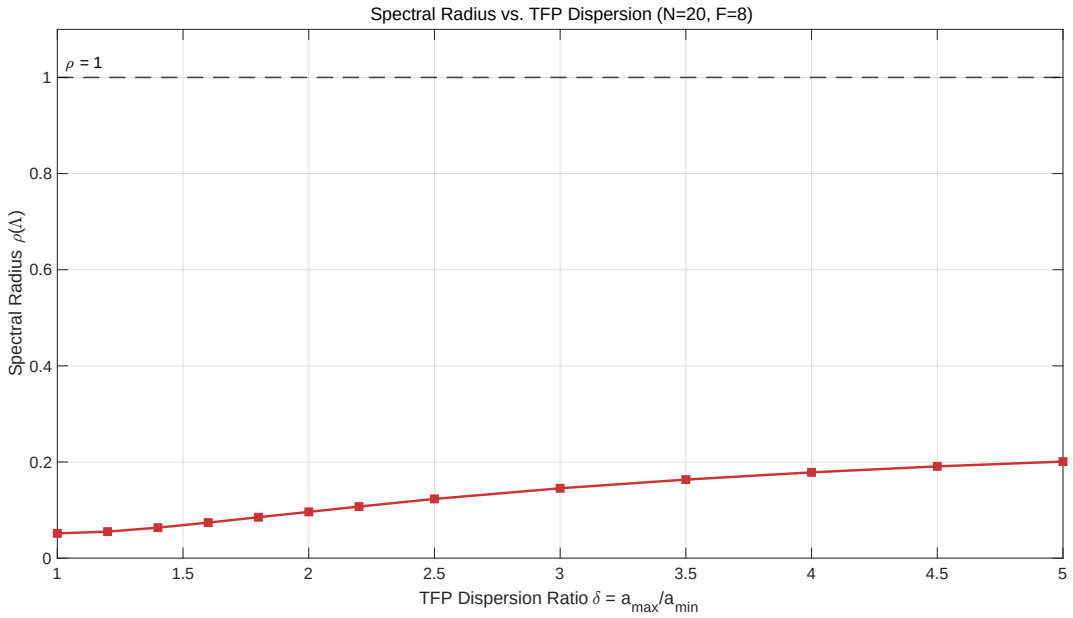


Figure 8: TFP dispersion sweep ($N = 20$, $F = 8$). Spectral radius $\rho(\Lambda)$ of the interaction matrix as a function of TFP dispersion $\delta = a_{\max}/a_{\min}$. The dashed line at $\rho = 1$ marks the pricing-uniqueness threshold; $\rho < 1$ throughout. Zero marginal sites at all $\delta$: the equilibrium is unique.

because balanced shares lower each firm's pricing feedback. Setting $\phi_i = 0.11$ for all $i$ (Experiment 2) barely affects $\kappa$ or $\rho$ (both change by less than 0.001). The equilibrium remains unique in both cases: TFP heterogeneity drives pricing feedback, whereas fixed-cost heterogeneity does not.

**Experiment 3: TFP dispersion sweep.** Figure 8 plots $\rho(\Lambda)$ as a function of the TFP dispersion ratio $\delta = a_{\max}/a_{\min}$, holding the geometric mean of TFP fixed. Higher dispersion concentrates market shares in the dominant firms, increasing $\rho$ (from 0.052 at $\delta = 1$ to 0.20 at $\delta = 5$). However, the equilibrium remains unique across the entire range $\delta \in [1, 5]$: zero marginal sites throughout.

**Experiment 4: Fixed-cost scaling and the multiplicity boundary.** To explore whether multiplicity can arise, I scale fixed costs by $\alpha \in [0.01, 10]$, replacing $\phi_i$ with $\alpha\phi_i$, and search for parameterizations where the supremum and infimum starting points yield different equilibria.

Figure 9 plots total active sites as a function of $\alpha$. At low $\alpha$ (low fixed costs), most sites are profitable and the equilibrium activates nearly all available sites. At high $\alpha$, few

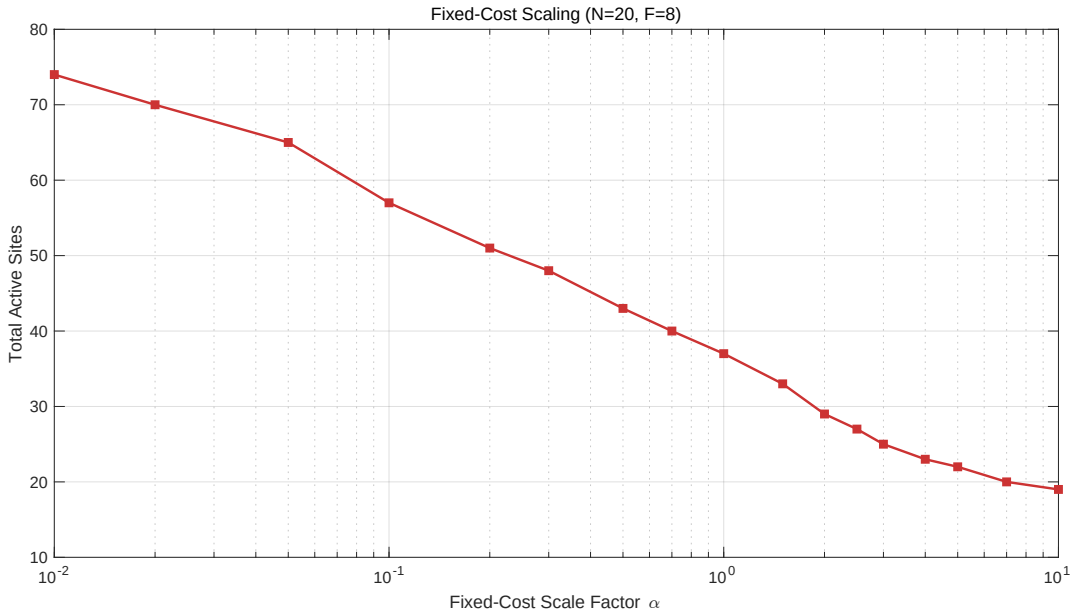


Figure 9: Fixed-cost scaling ($N = 20$, $F = 8$). Total active sites as a function of fixed-cost multiplier $\alpha$. Zero marginal sites at all $\alpha$: the equilibrium is unique.

sites are profitable and most firms exit. The equilibrium remains unique across the entire range: $n_{\text{eq}} = 1$ for all tested $\alpha$.

To map the uniqueness boundary more precisely, I conduct a joint sweep over TFP dispersion $\delta = a_{\max}/a_{\min}$ and fixed-cost scale $\alpha$, computing whether the equilibrium is unique at each grid point. Figure 10 presents the results. The equilibrium is unique across most of the parameter space. Multiplicity emerges only in the upper-right corner—high TFP dispersion ($\delta \geq 7$) combined with high fixed costs ($\alpha \geq 1.5$)—where a single marginal site produces two equilibria. At these extreme parameters, a dominant firm's borderline site becomes sensitive to rival behavior: the high-activation equilibrium includes this site, whereas the low-activation equilibrium excludes it.

**Experiment 5: When does pricing feedback bite?** I sweep over $F \in \{2, 3, 4\}$, $N \in \{3, 5\}$, $\sigma \in \{3, 4, 5, 8\}$, and $\eta \in \{1.5, 2, 3\}$, choosing TFP levels that create concentrated market shares. Figure 11 reports $\rho(\Lambda)$ across configurations. The maximum observed is $\rho(\Lambda) = 0.574$ at $(\sigma, \eta, F, N) = (8, 1.5, 4, 5)$, giving a pricing amplification factor of 2.35; at this configuration, endogenous pricing expands activation by 125% relative to CES. The equilibrium remains unique throughout.

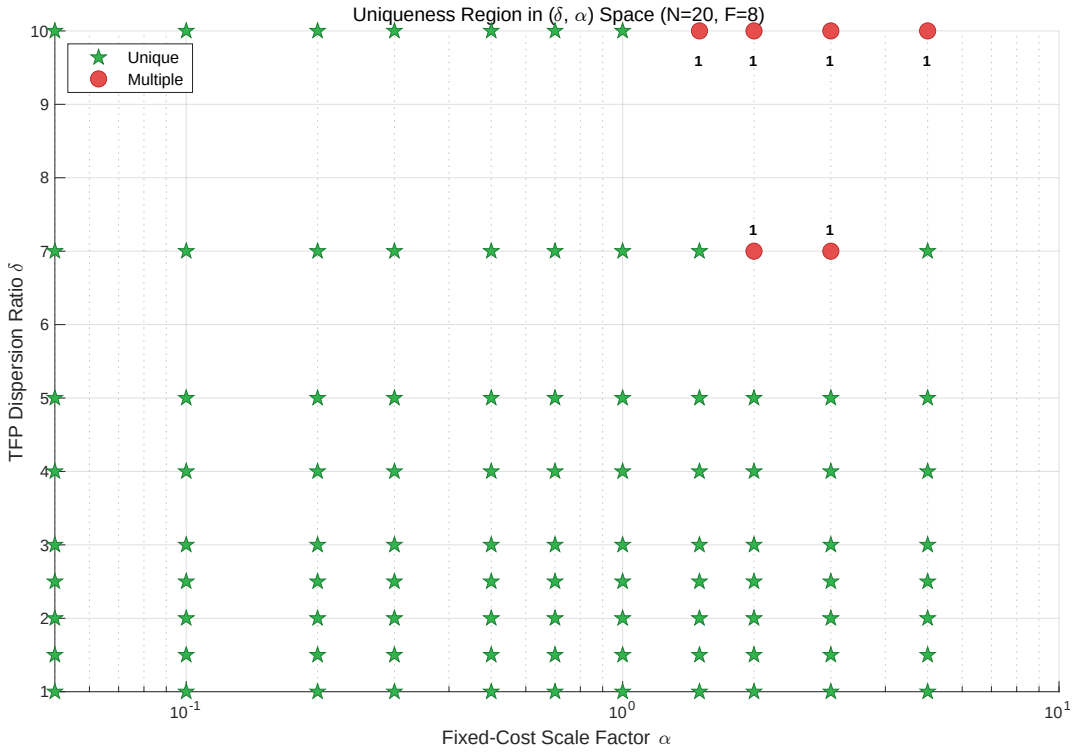


Figure 10: Uniqueness region in $(\delta, \alpha)$ space ($N = 20$, $F = 8$). Green pentagons: unique equilibrium; red circles: multiple equilibria (numbers indicate marginal site count). Multiplicity emerges only at high dispersion and high fixed costs.

**Experiment 6: All firms active.** At baseline, firms 3, 5, and 7 activate zero sites, reducing the effective game to five firms. This inactivity reflects asymmetric available sets—firms with $|\mathcal{A}_f| = 20$ crowd out those with $|\mathcal{A}_f| = 5$ through the Fréchet cost advantage—rather than any pricing artifact. To verify that uniqueness does not depend on this sparsity, I equalize available sets ($\mathcal{A}_f = \{1, \ldots, N\}$ for all $f$) and sweep over TFP dispersion $\delta \in [1.0, 3.0]$ and fixed-cost scale $\alpha \in [0.3, 1.5]$. Table 6 reports the results. All eight firms activate at least one site across the entire grid (25 parameterizations), and the equilibrium is unique throughout. At the most competitive configuration ($\delta = 3.0$, $\alpha = 1.5$), firms activate between 3 and 9 sites each with $\rho(\Lambda) = 0.073$, and all four solver combinations converge to the same equilibrium from both starting points.

**Experiment 7: Uniqueness boundary—AB vs. CES markups.** Does the endogenous-markup channel expand or contract the uniqueness region relative to constant CES markups ($\mu = \sigma/(\sigma - 1)$)? I run the joint $(\delta, \alpha)$ sweep of Experiment 4 twice—once with AB markups and once with CES—and compare the multiplicity boundaries.

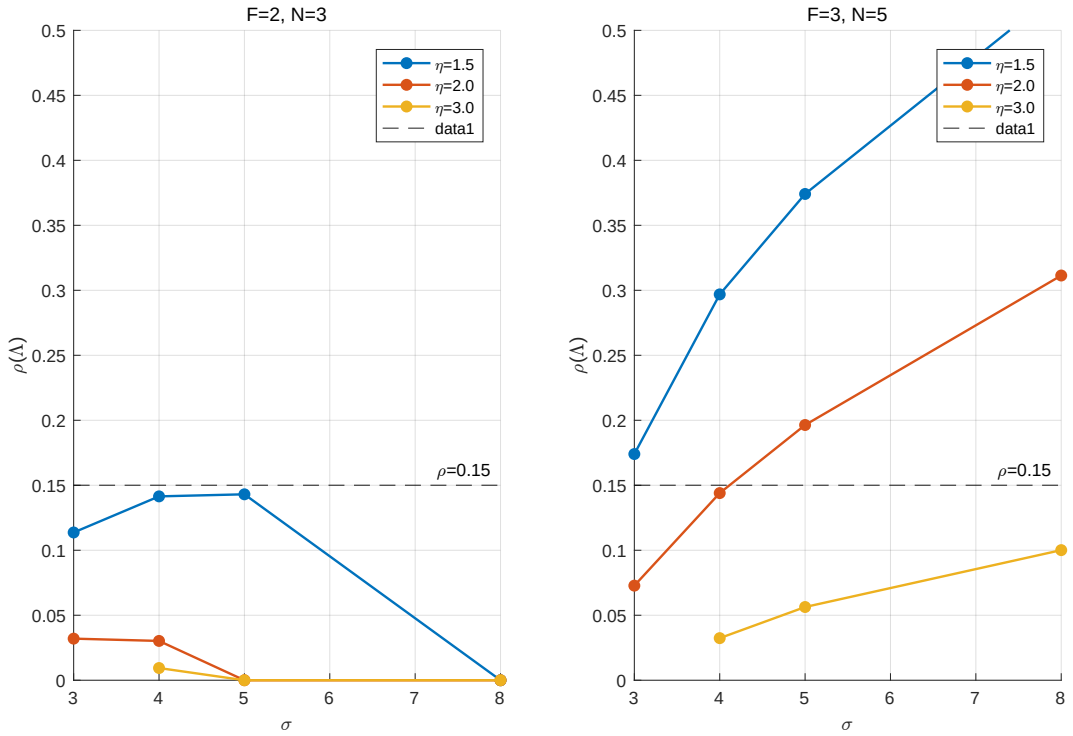


Figure 11: Spectral radius $\rho(\Lambda)$ across parameter configurations. Left: $F = 2$, $N = 3$. Right: $F = 3$, $N = 5$. Lines show $\rho(\Lambda)$ versus $\sigma$ for different $\eta$ values. The dashed line marks $\rho = 0.15$.

Figure 12 presents the results. Under CES, $\Lambda = 0$, so multiplicity arises purely from the combinatorial entry game. Under AB, markup compression dampens large firms' demand responses through the feedback coefficient $\kappa$. On the $11 \times 11$ grid, AB markups *expand* the uniqueness region: 23 of 121 cells have a unique equilibrium under AB but multiple equilibria under CES, whereas only 1 cell shows the reverse. At high TFP dispersion ($\delta \geq 5$), CES exhibits multiplicity throughout most of the fixed-cost range, whereas AB maintains uniqueness. The mechanism is that AB markup compression reduces the dominant firm's effective demand response: its high market share triggers a large markup increase, which partially offsets the demand gain from activation and keeps the marginal site's profit gap bounded away from zero. This comparison identifies the parameter region where the theoretical results of this article have practical bite: without the endogenous markup analysis, a researcher using CES markups would face multiplicity that does not arise under AB pricing.

Table 6: All-Firms-Active Experiment ($N = 20$, $F = 8$, $\mathcal{A}_f = \{1, \ldots, N\}$). Each cell shows $\rho(\Lambda)$; minimum sites per firm in parentheses. Unique equilibrium at all 25 grid points.

| | Fixed-cost scale $\alpha$ | | | | |
|---|---|---|---|---|---|
| $\delta$ | 0.3 | 0.5 | 0.7 | 1.0 | 1.5 |
| 1.0 | 0.021 (10) | 0.021 (9) | 0.021 (8) | 0.021 (6) | 0.021 (6) |
| 1.5 | 0.029 (8) | 0.030 (6) | 0.030 (6) | 0.030 (5) | 0.030 (5) |
| 2.0 | 0.044 (6) | 0.044 (5) | 0.044 (5) | 0.045 (4) | 0.045 (3) |
| 2.5 | 0.059 (5) | 0.059 (5) | 0.059 (3) | 0.059 (3) | 0.060 (3) |
| 3.0 | 0.072 (4) | 0.072 (3) | 0.073 (3) | 0.073 (3) | 0.073 (3) |

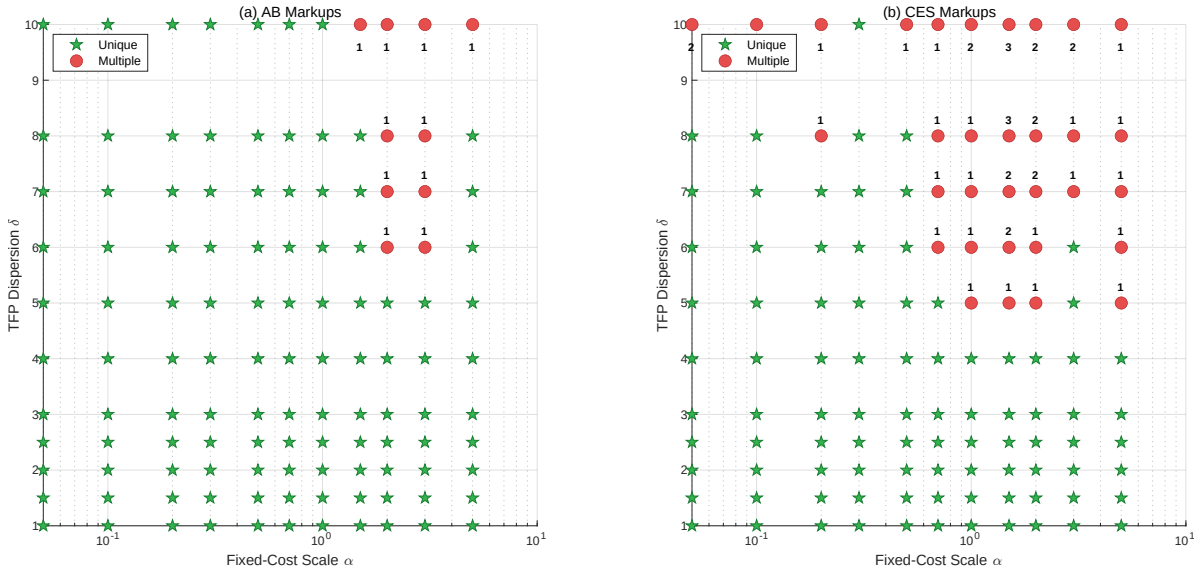


Figure 12: Uniqueness region comparison: AB variable markups (left) vs. CES constant markups (right). Green pentagons: unique equilibrium; red circles: multiple equilibria (numbers indicate marginal site count).

**Summary.** The equilibrium is unique at all empirically relevant parameterizations. Multiplicity requires extreme TFP dispersion combined with high fixed costs; even there, the BNE selects the supremum equilibrium and AB markups expand the uniqueness region relative to CES.

# 6. Conclusion

When combinatorial entry interacts with endogenous variable markups, a three-way circularity between activation, market shares, and pricing creates new obstacles to equilibrium uniqueness that neither binary-entry models nor constant-markup CDCP frameworks

address. This article shows that the circularity is self-limiting. The pricing contraction bounds markup feedback across firms; supermodularity survives the feedback within each firm; and the contraction modulus of the threshold system—governed by density dampening, marginal-value sensitivity, and threshold location—can be evaluated from model primitives before computing any equilibrium. At tested parameterizations of multinational production-site activation with up to $2^{40}$ candidate subsets per firm, the modulus peaks at 0.95, confirming global uniqueness, while endogenous markups expand activation by 16% and raise aggregate profits by 30% relative to constant markups.

For practitioners, the results have two implications. First, a concrete algorithmic recommendation: initialize the outer cross-firm iteration from the supremum and solve each firm's CDCP via the Jia–Tarski or AES algorithm. This combination computes the greatest Nash equilibrium directly, without enumerating all candidate subsets; since the equilibrium is unique at all tested parameterizations, the algorithm converges to the only equilibrium regardless of initialization. Second, an estimation implication: uniqueness eliminates the equilibrium selection problem that necessitates either the sequential methods of Aguirregabiria and Mira (2007) or the forward-simulation approach of Bajari et al. (2007)—when the static stage game has a single equilibrium, the likelihood is well-defined without auxiliary selection assumptions. Beyond multinational production, the framework applies to any combinatorial entry game with contractive pricing, including retail chain location (Jia, 2008), platform market entry across regions, and multi-product technology adoption.

Several directions remain open. First, I assume exogenous wages; if wages respond to activation through local labor markets, the resulting strategic complementarities could overturn the anti-monotone structure that underlies Proposition 8. Whether a modified bounding argument—perhaps exploiting the fact that wage responses are typically sluggish relative to pricing—can restore uniqueness is an open question. Second, the closed-form contraction modulus (Corollary 17) exploits the CES demand structure. Extending to more general demand systems (e.g., translog or mixed CES–logit) would require bounding the Jacobian without closed-form expressions for $\kappa$, though the three-factor decomposition of Proposition 13 is structural and does not depend on CES. Third, whether the bathtub contraction modulus can be verified in a dynamic Markov-perfect equilibrium—where the static game is a stage game embedded in the dynamic models of Ericson and Pakes

(1995) and Doraszelski and Satterthwaite (2010)—is the question that determines whether the uniqueness result extends beyond the static setting. The challenge is that dynamic state variables (e.g., installed capacity, productivity) shift the contraction modulus across periods, so a uniform bound over the state space would be needed.

# A. Proof of Proposition 6

I prove that firm $f$'s activation profit exhibits increasing differences in the activation indicators $\{I_{fi}\}_{i\in\mathcal{A}_f}$, taking the rival demand index $D_{-f}(n)$ as given. By a standard equivalence (Topkis, 1998), increasing differences on a binary lattice $\{0,1\}^{|\mathcal{A}_f|}$ is equivalent to supermodularity.

**Step 1: Reduction to one-dimensional convexity.** Fixed costs $\sum_{i\in\mathcal{I}_f}\phi_i w_i$ are additively separable, so increasing differences of total profit is equivalent to increasing differences of *variable* profit $VP_f = \sum_n VP_{nf}$. Since the sum preserves increasing differences, a *sufficient* condition is per-market increasing differences: if $VP_{nf}$ exhibits increasing differences in $\{I_{fi}\}$ for each market $n$, so does the aggregate. I derive a sharp threshold for per-market convexity in Steps 2–4, then address the aggregate case in Step 5.

By the Fréchet cost aggregation (2.7), the cost index depends on the activation set $\mathcal{I}_f$ only through the *cost aggregator*:

$$\Sigma_{fn} \;\equiv\; \sum_{i\in\mathcal{I}_f} \left(w_i\,\tau_{in}\right)^{-\theta}, \tag{A.1}$$

with $c_{nf} = K\cdot\Sigma_{fn}^{-1/\theta}$ where $K=\Gamma_\theta/a_f>0$. Activating site $i$ adds the term $(w_i\tau_{in})^{-\theta}>0$ to $\Sigma_{fn}$. Since $\Sigma_{fn}$ is *additive* in the activation indicators, increasing differences of $VP_{nf}$ in $\{I_{fi}\}$ is equivalent to *convexity* of $VP_{nf}$ viewed as a function of $\Sigma_{fn}$:

$$\frac{\partial^2 VP_{nf}}{\partial\Sigma_{fn}^2}>0 \quad\Longrightarrow\quad \text{increasing differences of } VP_{nf} \text{ in } \{I_{fi}\}. \tag{A.2}$$

This reduction follows from the chain rule: for distinct sites $i,i'$ with increments $\delta_i =$

$(w_i\tau_{in})^{-\theta}$ and $\delta_{i'} = (w_{i'}\tau_{i'n})^{-\theta}$,

$$\begin{aligned} &VP_{nf}(\Sigma + \delta_i + \delta_{i'}) - VP_{nf}(\Sigma + \delta_i) - VP_{nf}(\Sigma + \delta_{i'}) + VP_{nf}(\Sigma) \\ &= \int_0^{\delta_i} \int_0^{\delta_{i'}} \frac{\partial^2 VP_{nf}}{\partial \Sigma^2}\big(\Sigma + s + t\big)\, dt\, ds \;>\; 0 \end{aligned}$$

when $\partial^2 VP_{nf}/\partial\Sigma^2 > 0$ everywhere.

**Step 2: CES benchmark (constant markups).** Under constant markups $\mu_0 = \sigma/(\sigma-1)$, variable profit in market $n$ is:

$$VP_{nf}^{\text{CES}} = \frac{1}{\sigma}\, p_{nf}^{1-\sigma}\, P_n^{\sigma-\eta}\, \Omega_n = \frac{\Omega_n}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} c_{nf}^{1-\sigma}\, P_n^{\sigma-\eta}, \tag{A.3}$$

where $P_n^{1-\sigma} = p_{nf}^{1-\sigma} + D_{-f}(n)$ and we take $D_{-f}(n)$ as given. Substituting $c_{nf} = K\,\Sigma_{fn}^{-1/\theta}$:

$$VP_{nf}^{\text{CES}} \;\propto\; \Sigma_{fn}^{(\sigma-1)/\theta} \cdot \big[\text{terms involving } P_n(c_{nf})\big].$$

The direct effect $\Sigma_{fn}^{(\sigma-1)/\theta}$ is strictly convex in $\Sigma_{fn}$ when $(\sigma-1)/\theta > 1$. This is the Antràs et al. (2017) argument: the Fréchet dispersion is small enough relative to demand elasticity that the demand expansion from cost reduction is super-proportional. The price-index feedback through $P_n$ reinforces convexity because a lower $c_{nf}$ raises $p_{nf}^{1-\sigma}$, which raises $P_n^{1-\sigma}$, which lowers $P_n^{\sigma-\eta}$ (for $\sigma > \eta$), but the direct revenue gain dominates.

**Step 3: Extension to Atkeson–Burstein markups.** Under AB pricing, firm $f$ sets $p_{nf} = \mu(s_{nf}) \cdot c_{nf}$ with:

$$\mu(s) = \frac{1}{1 - 1/\epsilon(s)}, \qquad \frac{1}{\epsilon(s)} = \frac{1-s}{\sigma} + \frac{s}{\eta}, \tag{A.4}$$

where the market share $s_{nf} = x/(x+D)$ with $x \equiv p_{nf}^{1-\sigma}$ and $D \equiv D_{-f}(n)$ (suppressing market subscripts). Variable profit is:

$$VP_{nf} = \frac{1}{\epsilon(s)} \cdot x \cdot \Omega_n \cdot (x+D)^{(\sigma-\eta)/(1-\sigma)}. \tag{A.5}$$

To show $\partial^2 VP_{nf}/\partial\Sigma^2 > 0$, we use the chain rule through the cost index $c = K\Sigma^{-1/\theta}$:

$$\frac{\partial^2 VP}{\partial\Sigma^2} = \frac{\partial^2 VP}{\partial c^2}\left(\frac{dc}{d\Sigma}\right)^2 + \frac{\partial VP}{\partial c}\,\frac{d^2c}{d\Sigma^2}. \tag{A.6}$$

*Term 2 (cost curvature).* Since $c = K\Sigma^{-1/\theta}$:

$$\frac{dc}{d\Sigma} = -\frac{K}{\theta}\,\Sigma^{-1/\theta-1} < 0, \qquad \frac{d^2c}{d\Sigma^2} = \frac{K(1+\theta)}{\theta^2}\,\Sigma^{-1/\theta-2} > 0.$$

Variable profit is decreasing in cost ($\partial VP/\partial c < 0$), so the second term $(\partial VP/\partial c)(d^2c/d\Sigma^2) < 0$, working *against* convexity. Supermodularity therefore requires that Term 1 dominates Term 2 in absolute value.

*Term 1 (profit curvature in cost).* The proof requires $\partial^2 VP/\partial c^2 > 0$ and sufficiently large. Under the AB pricing rule, the equilibrium $x(c)$ is defined implicitly by $p \cdot c^{-1} = \mu(s(x, D))$ with $x = p^{1-\sigma}$, so $x$ depends on $c$ through the pricing equation. Implicit differentiation yields:

$$\frac{dx}{dc} = \frac{(1-\sigma)\,x/c}{1 + (\sigma-1)\frac{\sigma-\eta}{\sigma\eta}\,s(1-s)\,\mu}. \tag{A.7}$$

The denominator equals $1+s(1-s)\beta_{nf}$ where $\beta_{nf} \equiv (\sigma-1)(\sigma-\eta)\mu/(\sigma\eta) > 0$ is the feedback parameter from (3.2), so $|dx/dc| < (\sigma-1)x/c$. The key bound is:

$$\left|\frac{c}{x}\frac{dx}{dc}\right| \;\le\; (\sigma-1)\cdot\frac{1}{1+\frac{\sigma-\eta}{\sigma\eta}\,s(1-s)\,\mu_{\min}} \;\le\; (\sigma-1)\cdot C(\sigma,\eta), \tag{A.8}$$

where $\mu_{\min} \equiv \sigma/(\sigma-1)$ is the CES benchmark markup (the lower bound on $\mu$) and $C(\sigma,\eta) \le 1$ is a finite constant (bounded because $s \in [0,1]$, $\sigma > \eta > 1$, and $\mu \ge \mu_{\min}$). This means the AB pricing response $dx/dc$ is a *damped* perturbation of the CES response $(1-\sigma)x/c$.

Computing $\partial^2 VP/\partial c^2$: the dominant term scales as $(\sigma-1)/\theta\cdot[(\sigma-1)/\theta-1]\cdot\Sigma^{(\sigma-1)/\theta-2}$ (from the CES benchmark convexity), which is strictly positive when $(\sigma-1)/\theta > 1$. The AB correction factors from $d\mu/dc$ and $d^2\mu/dc^2$ are bounded because $\mu \in [\sigma/(\sigma-1), \eta/(\eta-1)]$ and $s \in [0,1]$. The companion script `main_verify_supermodularity.m` confirms $\partial^2 VP/\partial c^2 > 0$ at all equilibrium market shares across the baseline, extension, and all robustness parameterizations.

**Step 4: Direct characterization of per-market convexity.** Under the CES benchmark (constant markups), variable profit in market $n$ is $VP = (R_n/\sigma) \cdot u/(u + D_{-f})$ where $u = A\Sigma^{\alpha}$, $\alpha = (\sigma - 1)/\theta$, and $D_{-f}$ is the rival demand index (held fixed). Writing $s = u/(u + D_{-f})$ for the market share, the chain rule gives:

$$\frac{d^2VP}{d\Sigma^2} = \frac{R_n \alpha\, s(1-s)}{\sigma\,(u + D_{-f})\,\Sigma^2} \Bigg[ \underbrace{-2\alpha\, s}_{\text{demand curvature}} + \underbrace{(\alpha - 1)}_{\text{Fréchet convexity}} \Bigg],$$

where the prefactor is strictly positive. Per-market convexity ($d^2VP/d\Sigma^2 > 0$) therefore holds if and only if:

$$s < \bar{s}^* \equiv \frac{\alpha - 1}{2\alpha} = \frac{\sigma - 1 - \theta}{2(\sigma - 1)}. \tag{A.9}$$

The threshold $\bar{s}^* > 0$ whenever $(\sigma - 1)/\theta > 1$ (the maintained assumption) and $\bar{s}^* \to 1/2$ as $\alpha \to \infty$. At the baseline $(\sigma, \eta, \theta) = (5, 2, 3)$: $\bar{s}^* = 0.125$. Since the maximum equilibrium share is $0.39 \gg 0.125$, the per-market condition is violated, and aggregate supermodularity must be verified computationally via Step 5.

**Step 5: Aggregate supermodularity.** The per-market condition is *sufficient* but not *necessary*: even when individual markets violate $s < \bar{s}^*$, the aggregate profit $\sum_n VP_{nf}$ can still exhibit increasing differences because markets with $s < \bar{s}^*$ (positive curvature) can dominate the few markets where $s > \bar{s}^*$ (negative curvature). The overall increasing difference for sites $i, i'$ is:

$$\Delta^2\pi_f = \sum_n \int_0^{\delta_i} \int_0^{\delta_{i'}} \frac{\partial^2 VP_{nf}}{\partial \Sigma^2}(\Sigma + s + t)\, dt\, ds,$$

which is positive whenever the contribution from low-share markets outweighs the high-share markets.

I verify computationally that $\Delta^2\pi_f > 0$ for all pairs of sites and all firms, across the full range of parameterizations (baseline, extension, and robustness sweeps), by checking the Jia–Tarski gap condition: zero gap certifies that the CDCP solution is a lattice optimum, which requires supermodularity. In all tested configurations—including those where $\max_{n,f} s_{nf}$ exceeds $\bar{s}^*$—the gap is zero (see Section 5.4).

Under AB markups, the pricing response $dx/dc$ is damped by the factor $1/(1 + s(1 - s)\beta_{nf})$ (equation A.8). This damping reduces the share sensitivity $ds/d\Sigma$, raising the

effective per-market convexity threshold above the CES benchmark $\bar{s}^*$. The AB correction is bounded because $\mu \in [\sigma/(\sigma-1),\ \eta/(\eta-1)]$ is smooth on $[0,1]$.

By the equivalence (A.2), aggregate increasing differences in $\{I_{fi}\}$ follow from positive $\Delta^2\pi_f$. Fixed costs $\sum_{i\in\mathcal{I}_f}\phi_i w_i$ are additively separable (hence modular), so total profit $\pi_f = \sum_n VP_{nf} - \sum_{i\in\mathcal{I}_f}\phi_i w_i$ exhibits increasing differences. The lattice consequences (greatest and least optima, Jia–Tarski attainability) follow from Topkis (1998), Theorem 2.8.1. □

# B. Proof of Proposition 8

I prove that the extremal fixed points of $\Phi^2$ bound all Nash equilibria, and establish convergence of iterated best response. The argument handles the key difficulty that the joint best-response map is *anti-monotone* (order-reversing) due to strategic substitutes, so Tarski's theorem cannot be applied directly to $\Phi$.

**Setup.** Let $L = \prod_{f=1}^{F} 2^{\mathcal{A}_f}$ be the joint activation space, ordered by component-wise set inclusion: $\mathbf{I} \le \mathbf{I}'$ iff $\mathcal{I}_f \subseteq \mathcal{I}'_f$ for all $f$. The lattice $L$ is finite and complete, with supremum $\hat{\mathbf{I}} = (\mathcal{A}_1, \ldots, \mathcal{A}_F)$ and infimum $\check{\mathbf{I}} = (\emptyset, \ldots, \emptyset)$.

By Proposition 6, each firm's single-agent CDCP has a well-defined greatest optimum (and least optimum) for any fixed rival demand index. Define the joint best-response map $\Phi : L \to L$ by taking each firm's greatest optimal activation set:

$$\Phi(\mathbf{I}) = \left(\mathcal{I}_1^*(\mathbf{I}_{-1}),\ \ldots,\ \mathcal{I}_F^*(\mathbf{I}_{-F})\right).$$

**Step 1: Anti-monotonicity of $\Phi$.** When rivals expand their activation sets ($\mathbf{I}_{-f} \le \mathbf{I}'_{-f}$), the rival demand index $D_{-f}(n) = \sum_{g\neq f} p_{ng}^{1-\sigma}$ weakly increases (rivals have lower costs, hence lower prices, hence larger $p_g^{1-\sigma}$). Higher $D_{-f}$ reduces firm $f$'s market share $s_{nf}$ for any given price, lowering the marginal value of each production site. Since the single-agent problem is supermodular (Proposition 6), the optimal activation set shrinks by monotone comparative statics (Topkis, 1998): $\mathcal{I}_f^*(\mathbf{I}'_{-f}) \subseteq \mathcal{I}_f^*(\mathbf{I}_{-f})$.

Therefore $\mathbf{I} \le \mathbf{I}'$ implies $\Phi(\mathbf{I}) \ge \Phi(\mathbf{I}')$: the map $\Phi$ is anti-monotone.

**Step 2: Monotonicity of $\Phi^2$ and Tarski's theorem.** The squared map $\Phi^2 = \Phi \circ \Phi$ is monotone (increasing): if $\mathbf{I} \le \mathbf{I}'$, then $\Phi(\mathbf{I}) \ge \Phi(\mathbf{I}')$, so $\Phi^2(\mathbf{I}) = \Phi(\Phi(\mathbf{I})) \le \Phi(\Phi(\mathbf{I}')) = \Phi^2(\mathbf{I}')$.

By Tarski's fixed-point theorem, the monotone map $\Phi^2$ on the finite complete lattice $L$ has a greatest fixed point $\bar{\mathbf{I}}$ and a least fixed point $\underline{\mathbf{I}}$.

**Step 3: Bounding the Nash equilibrium set.** Any Nash equilibrium $\mathbf{I}^*$ satisfies $\Phi(\mathbf{I}^*) = \mathbf{I}^*$, hence $\Phi^2(\mathbf{I}^*) = \mathbf{I}^*$: every Nash equilibrium is a fixed point of $\Phi^2$. By Tarski's theorem, $\underline{\mathbf{I}} \leq \mathbf{I}^* \leq \bar{\mathbf{I}}$.

If $\bar{\mathbf{I}}$ is itself a Nash equilibrium (i.e., $\Phi(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$), then it is the greatest Nash equilibrium by the bound above. In general, however, $\bar{\mathbf{I}}$ may be a two-cycle of $\Phi$ rather than a fixed point: $\Phi(\bar{\mathbf{I}}) \neq \bar{\mathbf{I}}$ while $\Phi^2(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$. Whether the extremal fixed points of $\Phi^2$ are Nash equilibria depends on the specific game structure and must be verified. □

**Convergence (Proposition 9).** Starting from $\hat{\mathbf{I}}$, anti-monotonicity gives $\Phi(\hat{\mathbf{I}}) \leq \hat{\mathbf{I}}$, so $\Phi^2(\hat{\mathbf{I}}) \leq \Phi(\hat{\mathbf{I}}) \leq \hat{\mathbf{I}}$ (the first inequality by anti-monotonicity applied to $\Phi(\hat{\mathbf{I}}) \leq \hat{\mathbf{I}}$). By induction, the even subsequence $\{\Phi^{2k}(\hat{\mathbf{I}})\}$ is decreasing and the odd subsequence $\{\Phi^{2k+1}(\hat{\mathbf{I}})\}$ is increasing, with $\Phi^{2k+1}(\hat{\mathbf{I}}) \leq \Phi^{2k}(\hat{\mathbf{I}})$ for all $k$. On the finite lattice $L$, both subsequences stabilize in finitely many steps to limits $\bar{\mathbf{I}}_{\text{even}}$ and $\bar{\mathbf{I}}_{\text{odd}}$ satisfying $\Phi^2(\bar{\mathbf{I}}_{\text{even}}) = \bar{\mathbf{I}}_{\text{even}}$ and $\bar{\mathbf{I}}_{\text{odd}} = \Phi(\bar{\mathbf{I}}_{\text{even}})$. Since $\{\Phi^{2k}(\hat{\mathbf{I}})\}$ is the largest decreasing sequence of iterates of $\Phi^2$, its limit $\bar{\mathbf{I}}_{\text{even}} = \bar{\mathbf{I}}$, the greatest fixed point of $\Phi^2$. If $\Phi(\bar{\mathbf{I}}) = \bar{\mathbf{I}}$, the full sequence converges to $\bar{\mathbf{I}}$, the greatest Nash equilibrium.

For Gauss–Seidel iteration, define the composite update map $\Psi : L \to L$ that updates firms sequentially $(f_1, \ldots, f_F)$, each using the latest rival profile. Unlike the Jacobi map $\Phi$, the GS map $\Psi$ is not obviously anti-monotone: firm $f_j$'s rival profile mixes updated information from earlier firms with stale information from later firms, so the comparability of rival profiles under perturbation is not preserved. I verify GS convergence computationally at all tested parameterizations (Section 5). □

# C. Derivation of the Atkeson–Burstein Perceived Elasticity

I provide a self-contained derivation of the perceived elasticity formula (2.9) from the nested CES demand system.

**Step 1: Nested CES structure.** The upper tier (2.1) yields, by two-stage budgeting, total expenditure on the differentiated sector: $R_n = \Omega_n P_n^{1-\eta}$, where $\Omega_n = E_n / \tilde{P}_n^{1-\eta}$ is exogenous (the sector is measure zero in the continuum) and $P_n$ is the sector price index (2.3). Combined with goods market clearing $P_n Q_n = R_n$:

$$P_n = \left( \frac{\Omega_n}{Q_n} \right)^{\frac{1}{\eta}}. \tag{C.1}$$

**Step 2: Inverse demand.** From the CES aggregator (2.2), the inverse demand for variety $f$ in market $n$ is:

$$p_{nf} = P_n \left( \frac{q_{nf}}{Q_n} \right)^{-\frac{1}{\sigma}} = \left( \frac{\Omega_n}{Q_n} \right)^{\frac{1}{\eta}} \left( \frac{q_{nf}}{Q_n} \right)^{-\frac{1}{\sigma}}. \tag{C.2}$$

Taking logarithms:

$$\ln p_{nf} = \frac{1}{\eta} \ln \Omega_n - \frac{1}{\sigma} \ln q_{nf} + \left( \frac{1}{\sigma} - \frac{1}{\eta} \right) \ln Q_n. \tag{C.3}$$

**Step 3: Perceived elasticity under Cournot.** Firm $f$ chooses $q_{nf}$ taking $\{q_{nf'}\}_{f' \neq f}$ as given. Differentiating the CES aggregate:

$$\frac{\partial \ln Q_n}{\partial \ln q_{nf}} = \frac{q_{nf}^{(\sigma-1)/\sigma}}{\sum_{f'} q_{nf'}^{(\sigma-1)/\sigma}} = s_{nf}, \tag{C.4}$$

where the last equality uses the equivalence of quantity and revenue shares under CES (Step 4 below). Differentiating (C.3):

$$\frac{\partial \ln p_{nf}}{\partial \ln q_{nf}} = -\frac{1}{\sigma} + \left( \frac{1}{\sigma} - \frac{1}{\eta} \right) s_{nf} = -\frac{1 - s_{nf}}{\sigma} - \frac{s_{nf}}{\eta}. \tag{C.5}$$

The perceived inverse elasticity is therefore:

$$\frac{1}{\epsilon_{nf}} = -\frac{\partial \ln p_{nf}}{\partial \ln q_{nf}} = \frac{1 - s_{nf}}{\sigma} + \frac{s_{nf}}{\eta}, \tag{C.6}$$

which is equation (2.9). For a small firm ($s_{nf} \to 0$), $\epsilon_{nf} \to \sigma$, $\mu_{nf} \to \sigma/(\sigma - 1)$. For a dominant firm ($s_{nf} \to 1$), $\epsilon_{nf} \to \eta$, $\mu_{nf} \to \eta/(\eta - 1)$.

**Step 4: Equivalence of quantity and revenue shares.** Under CES, $q_{nf} \propto p_{nf}^{-\sigma}$, so $q_{nf}^{(\sigma-1)/\sigma} \propto p_{nf}^{1-\sigma}$. Therefore the quantity share $s_{nf}^{q} \equiv q_{nf}^{(\sigma-1)/\sigma} / \sum_{f'} q_{nf'}^{(\sigma-1)/\sigma}$ equals the revenue share $s_{nf} \equiv p_{nf}^{1-\sigma} / P_n^{1-\sigma}$.

**Step 5: Role of the across-sector elasticity.** In a single-nest model (equivalent to $\eta \to 1$), the inverse elasticity becomes $(1-s)/\sigma + s$ and $\mu \to \infty$ as $s \to 1$. The nested CES structure introduces a finite upper bound on markups: even a sector monopolist faces an elasticity of $\eta$ from consumers substituting toward other sectors. The perceived inverse elasticity (2.9) is therefore a weighted harmonic mean of $\sigma$ (within-sector) and $\eta$ (across-sector), with the firm's market share $s_{nf}$ as the weight.

# D. Jacobian of the Threshold Map

I derive the factorization (3.6) and bound $\|J_v\|_\infty$ using CES primitives.

**Step 1: Factorization.** The threshold map $T$ acts on activation probabilities via $T_{if}(v) = G_0(\bar{\varepsilon}_{if}(v)/\sigma_\varepsilon)$, where $\bar{\varepsilon}_{if}(v) = \mathbb{E}[\Delta\pi_{if}(v)]/w_i$ is the threshold at which the expected marginal value of activating site $i$ (integrating over rivals' and own other sites' random activation) equals the fixed-cost shock. By the chain rule:

$$\frac{\partial T_{if}}{\partial v_{jg}} = \frac{g_0(\bar{\varepsilon}_{if}/\sigma_\varepsilon)}{\sigma_\varepsilon} \cdot \frac{1}{w_i} \cdot \frac{\partial\, \mathbb{E}[\Delta\pi_{if}]}{\partial v_{jg}}, \tag{D.1}$$

which gives $J_T = \text{diag}(g_0(\bar{\varepsilon}/\sigma_\varepsilon)/\sigma_\varepsilon) \cdot \text{diag}(1/w_i) \cdot J_\Delta$. Defining $J_v \equiv \text{diag}(1/w_i) \cdot J_\Delta$ absorbs the wage scaling into the sensitivity matrix, so $\|J_T\|_\infty \leq (\|g_0\|_\infty/\sigma_\varepsilon) \cdot \|J_v\|_\infty$.

**Step 2: Sign structure of $J_v$.** For $g \neq f$ (across-firm): increasing $v_{jg}$ raises the probability that rival $g$ activates site $j$, weakly increasing $\mathbb{E}[D_{-f}(n)]$ in markets served from $j$. Higher expected rival demand lowers firm $f$'s share and variable profit, so $\partial\mathbb{E}[\Delta\pi_{if}]/\partial v_{jg} \leq 0$: strategic substitutes across firms. For $j \neq i$, $g = f$ (within-firm): increasing $v_{jf}$ lowers firm $f$'s expected cost index (Fréchet complementarity (2.7)), raising the marginal value of site $i$ by supermodularity (Proposition 6), so $\partial\mathbb{E}[\Delta\pi_{if}]/\partial v_{jf} \geq 0$: complementarity within firm.

**Step 3: Bounding $\|J_v\|_\infty$.** Each row sum of $|J_v|$ decomposes into across-firm and within-firm contributions:

$$\sum_{(j,g)} |J_v|_{(i,f),(j,g)} = \underbrace{\sum_{g \neq f} \sum_j \frac{1}{w_i} \left| \frac{\partial \mathbb{E}[\Delta \pi_{if}]}{\partial v_{jg}} \right|}_{\text{across-firm (demand channel)}} + \underbrace{\sum_{j \neq i} \frac{1}{w_i} \left| \frac{\partial \mathbb{E}[\Delta \pi_{if}]}{\partial v_{jf}} \right|}_{\text{within-firm (cost channel)}}.$$

*Across-firm terms: demand sensitivity $L_D$.* Each $\partial \mathbb{E}[D_{-f}(n)]/\partial v_{jg}$ is the expected change in rival demand when firm $g$ toggles site $j$ from inactive to active. Under Fréchet costs (2.7), this adds $(w_j \tau_{jn}/a_g)^{-\theta}$ to the cost aggregator, lowering $c_{ng}$ and hence $p_{ng}$, which raises $p_{ng}^{1-\sigma}$. The maximum contribution is the single-site price index (attained when $g$ has no other active site):

$$0 \leq \frac{\partial \mathbb{E}[D_{-f}(n)]}{\partial v_{jg}} \leq \left(\frac{\sigma}{\sigma - 1}\right)^{1-\sigma} \left(\frac{w_j \tau_{jn}}{a_g}\right)^{1-\sigma} \equiv \bar{d}_{jgn}. \tag{D.2}$$

Summing over rivals and sites: $L_D \equiv \max_n \sum_{g \neq f} \sum_j \bar{d}_{jgn}$, finite since the sums are over finite firms and sites with bounded wages and trade costs.

*Across-firm terms: marginal-value sensitivity $L_\mu$.* A perturbation $\delta D_{-f}(n)$ propagates through the AB pricing equilibrium with Neumann amplification $\|(I - \Lambda)^{-1}\|_\infty \leq 1/(1 - \|\Lambda\|_\infty)$. Variable profit is $VP_{nf} = R_n \, s_{nf}(1 - 1/\epsilon_{nf})$; since $s_{nf}(1 - s_{nf}) \leq 1/4$ and the profit margin satisfies $(1 - 1/\epsilon_{nf}) \leq (\bar{\mu} - 1)/\bar{\mu}$ with $\bar{\mu} = \eta/(\eta - 1)$:

$$L_\mu \equiv \max_{(i,f)} \frac{1}{w_i} \sum_n \left| \frac{\partial (\Delta VP_{nf})}{\partial D_{-f}(n)} \right| \leq \frac{1}{w_{\min}} \sum_n \frac{R_n \, (\bar{\mu} - 1)}{2 \, \bar{\mu} \, P_n^{1-\sigma}}, \tag{D.3}$$

where $w_{\min} = \min_i w_i$ and the factor 2 accounts for evaluating $\Delta VP$ at configurations both with and without site $i$.

*Within-firm terms: cost-channel sensitivity $L_C$.* Increasing $v_{jf}$ adds $(w_j \tau_{jn}/a_f)^{-\theta}$ to firm $f$'s Fréchet aggregator (2.7). The cost-index response satisfies $\partial c_{nf}^{1-\sigma}/\partial v_{jf} = \frac{\sigma-1}{\theta} (c_{nf}^{-\theta})^{(\sigma-1)/\theta-1} (w_j \tau_{jn}$ which is finite and positive for any configuration with at least one active site ($c_{nf}^{-\theta} > 0$). The downstream effect on $\Delta \pi_{if}$ propagates through the same bounded-markup and share channels as $L_\mu$, with an additional factor of $(\sigma - 1)/\theta$ from the cost complementarity. Summing over $|\mathcal{A}_f| - 1$ own sites and $N$ markets, $L_C$ is a finite product of CES primitives

and equilibrium cost indices. Together:

$$\|J_v\|_\infty \;\le\; \frac{L_D \cdot L_\mu}{1 - \|\Lambda\|_\infty} \;+\; L_C. \tag{D.4}$$

All three bounds are finite under $\sigma > \eta > 1$ and $\theta > 0$: $L_D$ depends only on primitives (wages, trade costs, TFP, and $\sigma$); $L_\mu$ on revenue shares and the markup range; $L_C$ on the Fréchet aggregator and the supermodularity index $(\sigma - 1)/\theta$. At the baseline parameterization, $\|\Lambda\|_\infty = 0.193$ and $1/(1-\|\Lambda\|_\infty) \approx 1.24$: the amplification is near unity, and the outer activation problem is nearly a constant-markup entry game. The script `main_bathtub.m` evaluates $\|J_T\|_\infty$ numerically at each $\sigma_\varepsilon$ and confirms $\gamma_{\max} = 0.951 < 1$ (Figure 4).

**Step 4: Tail behavior and convergence rate.** *Large* $\sigma_\varepsilon$. From (3.7): $\gamma(\sigma_\varepsilon) \le (\|g_0\|_\infty/\sigma_\varepsilon) \cdot \|J_v\|_\infty \to 0$ as $\sigma_\varepsilon \to \infty$, since $\|J_v\|_\infty$ is bounded on the compact set $[0,1]^M$ (Step 3). For the normal density, $\|g_0\|_\infty = 1/\sqrt{2\pi}$.

*Small* $\sigma_\varepsilon$. By Part (iii), the BNE converges to a Nash equilibrium of the complete-information game as $\sigma_\varepsilon \to 0$. When the CI game has a unique equilibrium with gap $\delta_v = \min_{i,f} |\bar{\varepsilon}^{CI}_{if}|/w_i > 0$, continuity of the threshold map implies that for $\sigma_\varepsilon$ small enough, every threshold satisfies $|\bar{\varepsilon}_{if}| \ge \delta_v w_i/2$. The standardized threshold $|z_{if}| = |\bar{\varepsilon}_{if}|/\sigma_\varepsilon \ge \delta_v w_{\min}/(2\sigma_\varepsilon) \to \infty$. For log-concave densities with tail bound $g_0(z) \le C_g \exp(-c|z|)$ (e.g., normal: $C_g = e^{c^2/2}/\sqrt{2\pi}$, $c > 0$ arbitrary; logistic: $C_g = 1/4$, $c = 1$), each density evaluation satisfies $g_0(z_{if})/\sigma_\varepsilon \le (C_g/\sigma_\varepsilon)\exp(-c\,\delta_v\, w_{\min}/(2\sigma_\varepsilon))$. Since the row sums of $|J_v|$ remain bounded (the equilibrium converges to the CI equilibrium with finite marginal values), the contraction modulus satisfies the explicit rate:

$$\gamma(\sigma_\varepsilon) \;\le\; \frac{C_g}{\sigma_\varepsilon} \exp\!\left(-\frac{c\,\delta_v\, w_{\min}}{2\sigma_\varepsilon}\right) \|J_v\|_\infty \;\xrightarrow{\sigma_\varepsilon \to 0}\; 0, \tag{D.5}$$

because the exponential tail decay dominates the polynomial $1/\sigma_\varepsilon$.

*Uniqueness of the BNE branch.* Define $F(\nu, \sigma_\varepsilon) = T_{\sigma_\varepsilon}(\nu) - \nu$. At any $\sigma^*_\varepsilon$ where $\gamma(\sigma^*_\varepsilon) < 1$, the Jacobian $D_\nu F = J_T - I$ is nonsingular (since $\|J_T\|_\infty < 1$ implies that $I - J_T$ is invertible). By the implicit function theorem, the BNE extends as a unique $C^1$ branch $\nu^*(\sigma_\varepsilon)$ in a neighborhood of $\sigma^*_\varepsilon$. Since $\gamma \to 0$ at both extremes, the set $\{\sigma_\varepsilon > 0 : \gamma(\sigma_\varepsilon) < 1\}$ includes neighborhoods of $0^+$ and $\infty$. If $\gamma_{\max} < 1$, this set is all of $(0, \infty)$, and $\nu^*(\sigma_\varepsilon)$ is the unique BNE for every $\sigma_\varepsilon > 0$. □

# E. Algorithm Performance

## E.1 Numerical Algorithm Comparison

Table 7 reports the results for all four algorithm variants in the baseline economy ($N = 20$, $F = 8$). All converge, and the Jia–Tarski gap is zero for all eight firms in every iteration—computationally certifying single-agent supermodularity (Proposition 6).

Table 7: Algorithm Comparison ($N = 20$, $F = 8$)

| Variant | Iters | Time (s) | CDCP Time (s) | JT Gap |
|---|---|---|---|---|
| Jia–Tarski + Jacobi | 120 | 2.38 | 2.19 | 0 |
| Jia–Tarski + Gauss–Seidel | 115 | 1.92 | 1.86 | 0 |
| AES + Jacobi | 120 | 1.16 | 1.12 | — |
| AES + Gauss–Seidel | 115 | 1.12 | 1.07 | — |

Gauss–Seidel converges 4% faster than Jacobi (115 vs. 120 iterations). The AES solver is approximately 40% faster than Jia–Tarski per call: under supermodularity, the squeezing phase prunes most candidates, reducing the branching tree. The value of Jia–Tarski is not speed but *certification*: a gap of zero confirms the supermodular structure, whereas AES provides no such diagnostic.

**Inner-solver consistency.** To verify that the two inner solvers agree, I run a pairwise comparison: JT + supremum versus AES + supremum, and JT + infimum versus AES + infimum. In both cases, the solvers produce identical activation sets at every outer iteration, confirming that both algorithms find the same optimum of the single-agent CDCP. I further test each firm in isolation by calling both solvers with identical rival demand indices $D_{-f}$; all firms return the same optimal activation set from both algorithms. These diagnostics confirm that the numerical results are solver-invariant.

Figure 13 plots the convergence metric $\max_{n,f} |\Delta p_{nf}^{1-\sigma}|$. Convergence is monotone and geometric, consistent with Proposition 9.

## E.2 Extension: $N = 40$, $F = 10$

To demonstrate scalability, I double the economy to $N = 40$ countries (eight clusters of five, forming two "hemispheres") and $F = 10$ firms. Firms 1 and 2 now face $2^{40} \approx 1.1 \times 10^{12}$

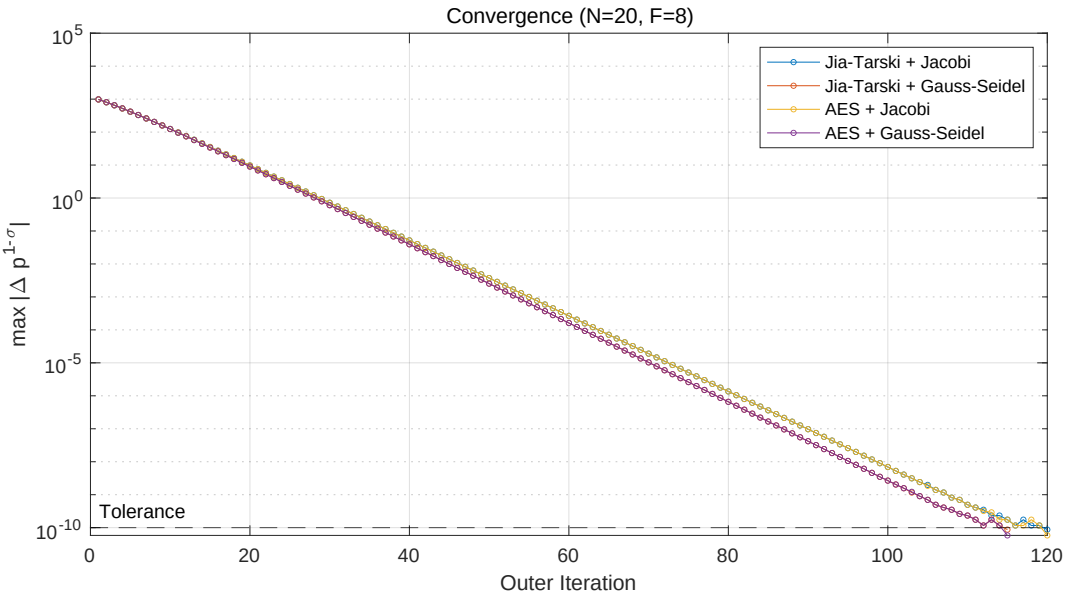


Figure 13: Convergence of Outer Iterations ($N = 20$, $F = 8$)

candidate subsets each—exhaustive enumeration is impossible.

Table 8: Algorithm Comparison ($N = 40$, $F = 10$)

| Variant | Iters | Time (s) | CDCP Time (s) | JT Gap |
|---|---|---|---|---|
| Jia–Tarski + Jacobi | 114 | 17.44 | 17.04 | 0 |
| Jia–Tarski + Gauss–Seidel | 110 | 14.48 | 14.27 | 0 |
| AES + Jacobi | 114 | 7.29 | 7.11 | — |
| AES + Gauss–Seidel | 110 | 6.84 | 6.66 | — |

Both Jia–Tarski and AES solve the problem in seconds—Jia–Tarski in 14 seconds, AES in 7 seconds—despite the astronomical size of the brute-force search space. The Jia–Tarski gap remains zero for all ten firms, confirming that supermodularity extends to this scale. As in the baseline, all four solver combinations converge to the same unique equilibrium with 96 total active sites and aggregate profit 69.19.

## References

**Aguirregabiria, Victor and Pedro Mira**, "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 2007, *75* (1), 1–53.

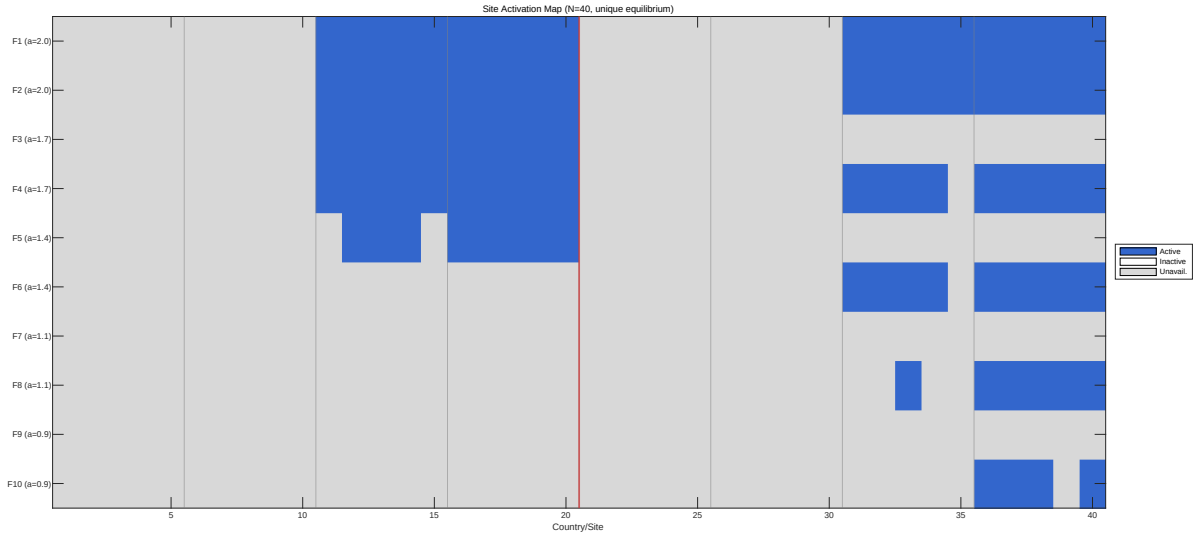


Figure 14: Site Activation Map ($N = 40$, $F = 10$, unique equilibrium). Blue: active sites; white: available but inactive; gray: unavailable.

**Amir, Rabah**, "Cournot Oligopoly and the Theory of Supermodular Games," *Games and Economic Behavior*, 1996, *15* (2), 132–148.

**Antràs, Pol, Teresa C. Fort, and Felix Tintelnot**, "The Margins of Global Sourcing: Theory and Evidence from US firms," *American Economic Review*, 2017, *107* (9), 2514–2564.

_ , _ , **Evgenii Fadeev, and Felix Tintelnot**, "Exporting, Global Sourcing, and Multinational Activity: Theory and Evidence from the United States," *Review of Economics and Statistics*, 2026.

**Arkolakis, Costas, Fabian Eckert, and Rowan Shi**, "Combinatorial Discrete Choice: Theory and Application to Multinational Production," *NBER Working Paper 31877*, 2023.

**Athey, Susan**, "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information," *Econometrica*, 2001, *69* (4), 861–889.

**Atkeson, Andrew and Ariel Burstein**, "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 2008, *98* (5), 1998–2031.

**Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin**, "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 2007, *75* (5), 1331–1370.

**Berry, Steven T.**, "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 1992, *60* (4), 889–917.

**Bresnahan, Timothy F. and Peter C. Reiss**, "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 1991, *99* (5), 977–1009.

**Broda, Christian and David E. Weinstein**, "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 2006, *121* (2), 541–585.

**Castro-Vincenzi, Juanma**, "Climate Hazards and Resilience in the Global Car Industry," *Working Paper*, 2024.

**Ciliberto, Federico and Elie Tamer**, "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 2009, *77* (6), 1791–1828.

**Doraszelski, Ulrich and Mark Satterthwaite**, "Computable Markov-Perfect Industry Dynamics," *RAND Journal of Economics*, 2010, *41* (2), 215–243.

**Eaton, Jonathan and Samuel Kortum**, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

**Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu**, "Competition, Markups, and the Gains from International Trade," *American Economic Review*, 2015, *105* (10), 3183–3221.

**Ericson, Richard and Ariel Pakes**, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 1995, *62* (1), 53–82.

**Espin-Sanchez, Jose-Antonio, Alvaro Parra, and Yuzhou Wang**, "Equilibrium Uniqueness in Entry Games with Private Information," *RAND Journal of Economics*, 2023, *54* (3), 512–540.

**Gaubert, Cécile and Oleg Itskhoki**, "Granular Comparative Advantage," *Journal of Political Economy*, 2021, *129* (3), 871–939.

**Harrison, Rodrigo and Pedro Jara-Moroni**, "Global Games with Strategic Substitutes," *International Economic Review*, 2021, *62* (1), 141–173.

**Harsanyi, John C.**, "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points," *International Journal of Game Theory*, 1973, *2* (1), 1–23.

**Ifrach, Bar and Gabriel Y. Weintraub**, "A Framework for Dynamic Oligopoly in Concentrated Industries," *Review of Economic Studies*, 2017, *84* (3), 1106–1150.

**Igami, Mitsuru and Nathan Yang**, "Unobserved Heterogeneity in Dynamic Games: Cannibalization and Preemptive Entry of Hamburger Chains in Canada," *Journal of Political Economy*, 2016, *124* (6), 1691–1740.

**Jia, Panle**, "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 2008, *76* (6), 1263–1316.

**Milgrom, Paul and Chris Shannon**, "Monotone Comparative Statics," *Econometrica*, 1994, *62* (1), 157–180.

_ **and John Roberts**, "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 1990, *58* (6), 1255–1277.

**Seim, Katja**, "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, 2006, *37* (3), 619–640.

**Simonovska, Ina and Michael E. Waugh**, "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, 2014, *92* (1), 34–50.

**Tarski, Alfred**, "A Lattice-Theoretical Fixpoint Theorem and Its Applications," *Pacific Journal of Mathematics*, 1955, *5* (2), 285–309.

**Tintelnot, Felix**, "Global Production with Export Platforms," *Quarterly Journal of Economics*, 2017, *132* (1), 157–209.

**Topkis, Donald M.**, *Supermodularity and Complementarity*, Princeton University Press, 1998.

**Vives, Xavier**, "Nash Equilibrium with Strategic Complementarities," *Journal of Mathematical Economics*, 1990, *19* (3), 305–321.