

Nash-in-Nash Equilibrium with Inefficient Contracts*

Diego Cussen[†] and Moeen Nehzati[‡]

First version: January 2025

This version: September 2026

Abstract

Inefficient contracts between firms and suppliers—contracts that do not maximize gains from trade—are ubiquitous. Research on bilateral oligopolies with inefficient contracts commonly employs the Nash-in-Nash bargaining equilibrium concept, but has been hampered by a lack of theoretical foundations for this equilibrium. We provide conditions for the existence and uniqueness of Nash-in-Nash equilibria in vertical relations games. These conditions imply that bargaining models are point identified, and can be verified in empirical work. We also provide an algorithm for fast equilibrium computation. These results advance our theoretical understanding of bilateral oligopolies and extend the toolset available to empirical researchers.

Keywords: Nash-in-Nash, bilateral oligopoly, bargaining, vertical relations.

JEL Classification: C78, D42, D43, L13.

*We are deeply indebted to our advisors, Giulia Brancaccio, Christopher Conlon, Dan Waldinger, and Alfred Galichon for their comments and advice. We also thank Joyee Deb, Dilip Abreu, Jess Benhabib, Alberto Bisin, Allan Collard-Wexler, Tommaso Denti, Andrés Elberg, Guillaume Fréchette, Gautam Gowrisankaran, Sinan Güntürk, Juan-Pablo Montero, Anh Nguyen, Álvaro Parra, David Pearce, Giulio Principi, Maher Said, Yassine Sbair Sassi, Nicolas Schutz (discussant), Paul Scott, Ennio Stacchetti, Audrey Tiew, Thomas Wollmann, and Ali Yurukoglu. Finally, we thank seminar and workshop participants at the International Industrial Organization Conference, New York University, and Pontificia Universidad Católica de Chile, and personnel of different industries who discussed evidence of the use of linear contracts. The views expressed in this paper, however, are ours and do not necessarily reflect the views of any industry personnel. All errors and omissions are our own.

See the replication package for our computation algorithm at https://github.com/MoeenNehzati/nash_in_nash_GD.

[†]Department of Economics, New York University (email: dc5004@nyu.edu). Corresponding author.

[‡]Department of Economics, New York University (email: sn3379@nyu.edu).

1 Introduction

Bargaining over input prices is central in relations between firms and their suppliers. Contracts shape profit functions and affect strategic interactions, so accounting for bargaining is key to studying issues such as bundling, horizontal mergers, vertical integration, price discrimination, exclusive contracts, and collusion.¹

Empirical industrial organization studies bilateral oligopolies to address these issues. Modeling bargaining in bilateral oligopolies is hard because negotiations between two firms affect other negotiations; beliefs about these externalities on and off the equilibrium path further complicate matters. To simplify the problem, Horn and Wolinsky (1988) developed the Nash-in-Nash equilibrium concept. In Nash-in-Nash, pairs of firms simultaneously negotiate à la Nash, assuming all other pairs sign non-renegotiable contracts. This framework has become a workhorse of the empirical literature on vertical relations.

Despite its central role in empirical research, neither existence nor uniqueness of a Nash-in-Nash equilibrium is ensured. This presents an econometric problem, because estimation assumes equilibrium conditions hold, and point identification often relies on uniqueness (Jovanovic, 1989; Galichon and Henry, 2011). Collard-Wexler et al. (2019) prove the existence and uniqueness of equilibrium in models where the gains from trade from negotiation are independent of the contracts firms sign. We call these *efficient* contracts, because they maximize gains from trade for a given bargaining network. This paper tackles a more general case: models where gains from trade can depend on the contracts all firms sign. We call these *inefficient* contracts. Efficient contracts are nested as a special case of our results.

To understand inefficient contracts, consider the case of linear (i.e., constant) input prices. These contracts are generally inefficient because they create double marginalization: redistributing surplus towards suppliers requires raising downstream firms' marginal costs, destroying industry profits (Crawford et al., 2018). To highlight the importance of inefficient contracts, Table 1 presents evidence that linear input prices are employed across several industries, countries, and years. Despite their inefficiency, it is clear that linear contracts are widespread.

This paper has three main contributions. First, we provide sufficient conditions for the existence of Nash-in-Nash equilibria for a wide class of vertical relations games. In these games, upstream suppliers and downstream retailers negotiate contracts à la Nash, following which retailers compete in a downstream market. Nash-in-Nash equilibria in these settings are solutions to systems of nonlinear equations, which generally do not exist. Our proof of existence ensures the criterion function for the estimation of bargaining models can be specified.

Existence does not rule out multiple, infinite, or even a continuum of equilibria. Thus, our second contribution is providing sufficient assumptions for local and global uniqueness of equilibria. Uniqueness matters because it supports the point identification of model parameters, and enables a clear interpretation of counterfactual exercises.

Our results do not depend on the number of firms in the market, the type of contracts firms negotiate (e.g., linear input prices, two-part tariffs, revenue sharing), or the number of inputs they trade. The results are also independent of the degree of contract inefficiency and the presence of vertical integration.

¹Bundling: Crawford and Yurukoglu (2012); horizontal mergers: Horn and Wolinsky (1988), Chipty and Snyder (1999), and Gowrisankaran et al. (2015); vertical integration: Crawford et al. (2018); price discrimination: Grennan (2013); exclusive contracts: Chambole and Molina (2023); collusion: Cussen and Montero (2024).

Table 1: Evidence of linear input prices

Product	Origin of data	Reference	Nash-in-Nash	Demand
Breakfast cereal	United States	Meza and Sudhir (2010)	Yes (TIOLI)	Mixed Logit
Cable TV	United States	Crawford and Yurukoglu (2012)	Yes	Mixed Logit
Cable TV	United States	Crawford et al. (2018)	Yes	Mixed Logit
Coffee	Chile	Noton and Elberg (2018)	Yes	Mixed Logit
Mortgage brokerage	United Kingdom	Robles-Garcia (2020)	Yes	Mixed Logit
Health services	Chile	Cuesta et al. (2025)	Yes	Mixed Logit
Pharmaceuticals	China	Barwick et al. (2025)	Yes	Mixed Logit
Cancer drugs	United States	Dix and Lensman (2025)	Yes	Mixed Logit
Coronary stents	United States	Grennan (2013, 2014)	Yes ($\times 2$)	RCNL
Beer	United States	Asker (2016), Miller and Weinberg (2017b)	Yes (TIOLI)	RCNL
Pharmaceuticals	Norway	Dubois and Sæthre (2020)	Yes	$\approx$ Nested Logit
Medical services	United States	Gowrisankaran et al. (2015)	Yes	Logit
Medical services	United States	Ho and Lee (2017)	Yes	Logit
Many (imports)	United States	Alviarez et al. (2023)	Yes	Iso-elastic
Eggs	Chile	Cussen and Montero (2024)	Yes	Linear
Concrete	Denmark	Albæk et al. (1997)	No	—
Refrigerated juice, tuna	United States	Kadiyali et al. (2000)	No	—
Gasoline	United States	Hastings and Gilbert (2005)	No	—
e-books	Several	Gilbert (2015)	No	—
Medical devices	United States	Grennan and Swanson (2020)	No	—
Bricks	United Kingdom	Beckert et al. (2020)	No	—
Carbonated beverages	United States	Luco and Marshall (2020)	No	—
Defined contribution plans	United States	Bhattacharya and Illanes (2022)	No	—

Notes: Each row is a different sample, and contains the product(s), paper that presents the data, whether the paper uses the Nash-in-Nash equilibrium concept and, if so, the demand system. Following Lee et al. (2021), we consider models with take-it-or-leave-it (TIOLI) offers as a limit case of Nash-in-Nash. RCNL is random coefficients nested logit. Ho and Lee (2017) use a variant of logit. The model in Dubois and Sæthre (2020) is similar to, but different from, nested logit. Gilbert (2015) presents evidence on Amazon’s sales of e-books in the US, but the input pricing model is the same when selling in other countries. For beer we consider two papers: Asker (2016) and Miller and Weinberg (2017b); Miller and Weinberg (2017a) present a model. Linear input prices may be used in some US states due to retail price maintenance agreements. In Norway, side-payments are forbidden (Brekke et al., 2015), so linear fees are used instead (Dubois and Sæthre, 2020).

To prove existence and uniqueness of a Nash-in-Nash equilibrium, our main assumptions are that there is a unique pure strategy equilibrium in the downstream competition game, that profit functions are smooth, and that firms negotiating à la Nash have a unique bargaining solution. Under these conditions we can use the Brouwer and Poincaré-Hopf theorems to prove the existence and (at least local) uniqueness of a Nash-in-Nash equilibrium for Lebesgue-almost every parameter vector in the parameter space.

This last condition means that if a model has infinite equilibria, then an arbitrarily small perturbation of its parameters yields a model with (at least locally) unique equilibria. Economic theory provides smoothness and existence and uniqueness of equilibrium results for the downstream competition games we study, so two of our main assumptions can be verified using models’ primitives.

Uniqueness of a bargaining solution is different. With efficient contracts quasi-concavity of the Nash

product (the bargaining problem’s objective function’s) is satisfied. With inefficient contracts, however, the Nash product can be a highly nonlinear function. This means one cannot generally prove the Nash product is quasi-concave using primitives.

To address this problem, we first note that the unique bargaining solution assumption is not necessary for either existence or uniqueness of a Nash-in-Nash equilibrium. Then we develop our third contribution: an algorithm that quickly computes Nash-in-Nash equilibria using gradient descent. This algorithm has four useful properties. First, it only fails to converge if the bargaining solution is not unique. Second, it is fast, allowing researchers to quickly verify uniqueness of the bargaining solution and explore the contract space for multiple equilibria. Third, our algorithm may recycle code used for estimation, meaning it is easy to implement once such code exists.

Fourth, the algorithm uses our existence and uniqueness conditions to deal efficiently with the two-stage structure of the games we study. Nash-in-Nash equilibrium computation is costly due to the multi-stage structure of vertical relations games. A solution used in empirical work is specifying and estimating one-stage games, where input price negotiations and downstream competition occur simultaneously. This reduces equilibrium computation costs substantially, but risks introducing bias of unknown size and direction in the estimation (Crawford et al., 2018). Our algorithm avoids these concerns. These properties of our algorithm are only ensured under the assumptions we use to prove existence and uniqueness of a Nash-in-Nash equilibrium.

As proof of concept, we use our algorithm to compute equilibria in Grennan (2013)’s model of the market for stents, and find that it is fast and tractable. It takes under 20 seconds to compute a Nash-in-Nash equilibrium in a grid of 10^{24} gridpoints. This speed allows us to explore the contract space for multiple equilibria, and we find repeated convergence to the same equilibrium in 100 runs using random starting points.

To ensure our results are applicable, we use extant and new results to prove that models with discrete choice demand and constant returns to scale can satisfy our assumptions for existence and local and global uniqueness. This shows that our assumptions are mutually consistent and that standard models used in IO are well suited for modeling vertical relations with Nash-in-Nash bargaining.

Related Literature. A significant amount of research has been devoted to modeling bargaining in bilateral oligopolies. Aside from Collard-Wexler et al. (2019), the papers closest to ours are Ho and Lee (2019), Rey and Vergé (2020), and Gross and Savelle (2025). Rey and Vergé (2020) prove the existence of Nash-in-Nash equilibria when gains from trade do not depend on contracts. Their setting differs from ours in that, in their model, contracts are private information. Gross and Savelle (2025) prove the existence of a symmetric Nash-in-Nash equilibrium in symmetric models where bargaining and competition are simultaneous. Our sequential setting, in contrast, allows for arbitrary asymmetries across firms.

Considering efficient contracts, Ho and Lee (2019) make the network of firms that bargain endogenous. Their equilibrium concept is called Nash-in-Nash with Threat of Replacement because network formation occurs via one side of the market unilaterally offering firms on the other side participation in a network, from which they may be excluded.

Nash-in-Nash with Threat of Replacement differs from Nash-in-Nash in that the equilibrium network of agreements is determined while firms bargain. In Nash-in-Nash, by contrast, the observed network (but not the off-path network) is fixed by the time firms negotiate. Modeling network formation is

beyond the scope of this paper, but previous work shows that Nash-in-Nash is consistent with the network being an equilibrium outcome. Lee and Fong (2013)’s network formation game provides a microfoundation for the network in Nash-in-Nash, and Rey and Vergé (2020) show that the network-formation stage can be modeled using coalition-proof Nash equilibria (Bernheim et al., 1987). Ide and Montero (2024) rationalize the observed network using exclusive contracts.

Our paper also relates to the applied literature on vertical relations with bargaining, with linear and nonlinear input prices. On the empirical side, these papers study rollups (Asil et al., 2026) and buyer power (Demirer and Rubens, 2025), among other issues. On the ex ante antitrust policy evaluation side, Sheu and Taragin (2021) and Panhans (2024) use simulations to study the competitive effects of vertical integration under various types of contracts. We contribute to both literatures by showing conditions under which equilibria in those models exist and are unique.

Finally, we contribute to the literature on equilibrium computation. Galichon (2016) and Galichon et al. (2025) use tools from optimal transport to prove computation results in various models, including transport networks, matching, and discrete choice. Several results depend on convexity or gross substitution, conditions that do not generalize to vertical relations games. Nguyen and Vohra (2018) use Scarf’s lemma as an alternative to deferred acceptance in matching models, where the value of relations does not depend on which other relations have been formed. We provide an algorithm that converges quickly and reliably to Nash-in-Nash equilibria when none of these conditions hold.

We proceed as follows. Section 2 reviews existing research that predicts the use of inefficient contracts and presents additional evidence of their widespread use. Section 3 characterizes the class of models we study. Section 4 introduces sufficient conditions for the existence and uniqueness of a solution to firms’ bargaining problems. Existence of a Nash-in-Nash equilibrium is proven in Section 5. Section 6 presents local and global uniqueness results. Section 7 presents our equilibrium computation algorithm. Section 8 shows how to verify that our sufficient conditions for existence and uniqueness hold, with a focus on discrete choice demand models. Section 9 provides examples and Section 10 concludes.

2 Inefficient Contracts

This section provides background on the use of inefficient contracts. First we review existing theoretical results that predict firms will sign inefficient contracts in various settings. Then we present further evidence of their widespread use.

2.1 Why Firms Use Inefficient Contracts

In theory, firms can solve the double marginalization by signing two-part tariff contracts, in which the buyer purchases inputs at marginal cost and compensates the supplier with a lump-sum transfer. This maximizes firms’ gains from trade, and bargaining is reduced to the size of the transfer.

In practice, firms may not sign efficient two-part tariff contracts for several reasons. First, arbitrage in the wholesale market may prevent suppliers from using two-part tariffs (Tirole, 1988). Second, Calzolari et al. (2020) show that two-part tariffs cause welfare losses if there are information frictions. Building on that model, Cussen and Montero (2022) show that, for high enough frictions, firms in bilateral oligopolies will choose linear prices over two-part tariffs.

Third, a vertically integrated firm that supplies its rivals may optimally constrain itself to linear input prices (Kourandi and Pinopoulos, 2022). Fourth, a multiproduct monopoly supplier that cannot commit to a unit price schedule may optimally negotiate linear prices with retailers to avoid dissipating downstream rents (Ide and Montero, 2024).

Finally, Iyer and Villas-Boas (2003) consider incomplete contracts that arise when product characteristics are not fully contractible. This incompleteness encourages opportunism, and the authors show that firms will avoid two-part tariffs in these cases.

These results explain why linear input prices are widespread. But nonlinear contracts, which include two-part tariffs, quantity-forcing contracts, buyback clauses and quantity discounts, are also observed (Cachon, 2003). Nonlinear contracts, however, can also be inefficient.

To begin with, risk aversion and information frictions may cause welfare losses proportional to the transfer in a two-part tariff (Calzolari et al., 2020). Firms may also use nonlinear contracts to exert market power. An example is slotting fees: lump-sum transfers from suppliers to retailers who sell their products. Theoretically, slotting fees may be used to extract surplus from consumers as well as suppliers (Caprice and von Schlippenbach, 2013), to facilitate collusion (Shaffer, 1991), or for foreclosure (Asker and Bar-Isaac, 2014). In the next subsection we review evidence that these mechanisms are at play.

2.2 Evidence of Inefficient Contracts

Table 1 presents evidence of the widespread use of linear input prices. More evidence of their use comes from the practices of large retail chains in the U.S., that historically negotiated linear prices with their suppliers (Elberg and Noton, 2024). Conversations with personnel of the fishing, agriculture and wine production industries have yielded additional evidence of *negotiated* linear input prices. Zheng (2024) reports that the Chinese government guided firms in the upstream coal market to negotiate linear prices.

Further evidence comes from antitrust cases. These include the antitrust cases of collusion in the thread industry in the European Union, and in fiberglass and urethane in the U.S.² Dobson and Waterson (1997) cite numerous reports from the United Kingdom’s Monopolies and Mergers Commission.

Moving on to nonlinear prices, several papers find evidence of inefficiency. In the markets Hristakeva (2022) studies, for example, firms use slotting fees. Yet, unit prices are consistently above suppliers’ marginal costs, and the non-negativity constraint of the slotting fees sometimes has a positive shadow value. This implies that efficiency is not the only rationale for those transfers. The author explains the result with the foreclosure mechanisms proposed by Asker and Bar-Isaac (2014).

Elberg and Noton (2024) use a uniquely detailed dataset of the supply contracts of a large retailer to study the characteristics of nonlinear contracts in general, and trade allowances (including slotting fees) in particular. They find evidence consistent with the market power view of nonlinear contracts, and that slotting fees are not being used to reduce double marginalization. Rao and Mahi (2003) also find evidence consistent with the market power view.

The takeaways from this section are: (1) theory predicts that inefficient contracts will be widely used, and (2) ample evidence supports these predictions. This does not imply that contracts are never efficient,

²Threads: Commission of the European Communities, 14.09.2005, Case COMP/38337/E1/PO/ Thread, 112, 159-60. Fiberglass: *Reserve Supply v. Owens-Corning Fiberglas* 971 F. 2d 37 (7th Cir., 1992). Urethane: *In Re: Urethane Antitrust Litigation*, No. 13-3215 (10th Cir., 2014).

but suggests the importance of having equilibrium existence and uniqueness results for bargaining over inefficient contracts.

3 Setting

This section has two parts. First, we define the class of models we study. These are vertical relations models where bilateral oligopolies of suppliers and retailers negotiate contracts. After that, retailers compete downstream. Then we discuss the equilibrium in the downstream competition stage.

3.1 Primitives

There are two types of players: upstream suppliers U_i , $i \in \{1, \dots, I\} \equiv \mathcal{I}$, and downstream retailers D_j , $j \in \{1, \dots, J\} \equiv \mathcal{J}$. The game has two stages. In the first stage a network of supplier-retailer pairs (“bargaining pairs”) simultaneously negotiate contracts à la Nash. At the start of the second stage all contracts are revealed, and retailers compete in a downstream market.

A bargaining pair is denoted (ij) , and the network of all pairs is $G \subset \mathcal{I} \times \mathcal{J}$. The contract firms in pair (ij) negotiate is a vector of K_{ij} contract terms $\mathbf{x}_{ij} \in X_{ij} \subset \mathbb{R}^{K_{ij}}$ (**bold** denotes a vector). The set of admissible contracts for (ij) , X_{ij} , is nondegenerate, compact, and convex. An element of $\mathbf{x}_{ij}$ (a “contract term”) is x_{ijk} , $k \in \{1, \dots, K_{ij}\} \equiv \mathcal{K}_{ij}$. D_j ’s bargaining power relative to U_i is $b_{ij} \in (0, 1)$. We call b_{ij} the bargaining parameter.

An example is a contract specifying linear input prices for K_{ij} inputs. Another is a two-part tariff specifying $K_{ij} - 1$ unit prices and a lump-sum transfer.

The vector of contract terms negotiated by all other bargaining pairs is $\mathbf{x}_{-(ij)} \in X_{-(ij)} \subset \mathbb{R}^{K_{-(ij)}}$. $X_{-(ij)}$ is compact and convex. During their negotiations D_j and U_i assume that all other bargaining pairs have agreed to an $\mathbf{x}_{-(ij)}$, and that no other bargaining pair will renegotiate its contract if bargaining between D_j and U_i breaks down.

Retailer D_j ’s strategy set in the second stage is the set $P_j \subset \mathbb{R}^{n_j}$. Thus, $P \equiv \prod_{j \in \mathcal{J}} P_j$.³ So, in our setting a model is a list of players $(\mathcal{I}, \mathcal{J})$, together with a list of expected profit functions $(\pi_f(\cdot; \cdot))_{f \in (\mathcal{I}, \mathcal{J})}$, parameterized by a finite-dimensional parameter vector $\boldsymbol{\theta} \in \Theta$, strategy sets $(P_f)_{f \in (\mathcal{I}, \mathcal{J})}$, a network of bargaining pairs $G \subset \mathcal{I} \times \mathcal{J}$, and admissible contract sets $(X_{ij})_{(ij) \in G}$. Fixing the sets of players, strategies, the network of bargaining pairs, admissible contracts sets, and profit functions up to $\boldsymbol{\theta}$, a model is defined by the vector of parameters $\boldsymbol{\theta}$.

Notation. The number of contract terms in the model is $K \equiv K_{ij} + K_{-(ij)}$. With some abuse, $X \equiv X_{ij} \times X_{-(ij)}$ and $\mathbf{x} \equiv (\mathbf{x}_{ij}, \mathbf{x}_{-(ij)})$ for all $(ij) \in G$. A vector of admissible contract terms is $\mathbf{x} \in X$.

Discussion. Our setting encompasses a large class of models used in IO. As Lee et al. (2021) show, the timing is consistent with firms taking other actions before, during, or after the bargaining stage that may affect bargaining outcomes. Without loss, the second stage can stand in for several stages.

We are agnostic about the type of competition in the second stage. It may be in prices or quantities, with or without product differentiation, and may include capacity constraints. Prices may be linear or nonlinear, and retailers may sell multiple products. We use $\mathbf{p}_j$ to denote a strategy vector for D_j

³Without loss of generality and some abuse, we let suppliers’ strategy sets be empty, as all they can do in the second stage is comply with the contracts they signed in the first stage.

because we will frequently apply our results in games where price is the strategic variable. Our results, however, also apply to games where firms choose quantities or other variables.

3.2 Equilibrium in the Second Stage

Firms' profits in the bargaining stage are their profits in the second stage as functions of contract terms $\mathbf{x}$ only. For those functions to exist there must be an equilibrium in the second stage.

Assumption 1. *For each $\mathbf{x} \in X$ (and for every (ij) , for every $\mathbf{x}_{-(ij)} \in X_{-(ij)}$), there exists a unique pure strategy equilibrium in the second stage. Firms' profits, on and off-path, can be written as functions of $\mathbf{x}$.*

Considerable research establishes the existence and uniqueness of equilibria for models used in IO. These include CES, logit, nested logit, and mixed logit, with single or multiproduct firms, and different cost functions, pricing structures, and degrees of substitution. See Caplin and Nalebuff (1991), Gallego et al. (2006), Liu (2006), Konovalov and Sándor (2010), Morrow and Skerlos (2010b), Aksoy-Pierson et al. (2013), Gallego and Wang (2014), Nocke and Schutz (2018, 2025), Garrido (2024), and Cussen (2026).⁴

4 The Bargaining Stage

Given existence of an equilibrium in the second stage, this section presents firms' bargaining problems in the first stage, and conditions that ensure the existence of a continuous bargaining solution.

4.1 The Bargaining Problem

Consider pair $(ij) \in G$, and let $f \in \{U_i, D_j\}$. Given $\mathbf{x}_{-(ij)}$, if U_i and D_j agree on contract $\mathbf{x}_{ij}$, f 's on-path profits are $\hat{\pi}_f(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)})$. If bargaining breaks down, its off-path profits are $\bar{\pi}_f(\mathbf{x}_{-(ij)})$. So, f 's gains from trade are

$$\Delta\pi_f(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}) \equiv \hat{\pi}_f(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}) - \bar{\pi}_f(\mathbf{x}_{-(ij)}).$$

Given $\mathbf{x}_{-(ij)}$, we define f 's participation constraint as the closure, relative to X_{ij} , of the contracts that give f strictly positive gains from trade. Thus, f 's participation constraint is the correspondence PC_{ij}^f from $X_{-(ij)}$ to X_{ij} given by

$$PC_{ij}^f(\mathbf{x}_{-(ij)}) := cl_{ij} \left(\left\{ \mathbf{x}_{ij} \in X_{ij} : \Delta\pi_f(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}) > 0 \right\} \right) \quad (1)$$

where cl_{ij} denotes closure relative to X_{ij} .⁵ Participation constraints depend on $\mathbf{x}_{-(ij)}$ because D_j 's residual demand generally depends on the prices or quantities of all retailers. The intersection of U_i and

⁴There is no proof of existence of an equilibrium in models with mixed logit demand under general specifications; the works cited here apply to a subset of models. Morrow and Skerlos (2010a) and Morrow (2015) provide general conditions for equilibrium candidates to exist. Cussen (2026) proves that a large class of models with linear utility and constant returns to scale satisfy those conditions. Nehzati and Cussen (2026) prove that the number of equilibrium candidates is finite for Lebesgue-almost every parameter vector under linear utility and constant returns to scale. Morrow and Skerlos (2011) provide a fast and accurate algorithm for computing equilibria once equilibrium candidates exist.

⁵Similarly, int_{ij} means interior relative to X_{ij} , as opposed to relative to $\mathbb{R}^{K_{ij}}$. Under Assumption 2 both interiors are equal (see proof of Lemma 1 in Appendix A).

D_j 's participation constraints is called the *bargaining set*, denoted $B_{ij}(\mathbf{x}_{-(ij)})$. Thus the bargaining set is the closure of the set of jointly profitable contracts. Jointly profitable contracts are contracts such that, given $\mathbf{x}_{-(ij)}$, both firms get positive gains from trade.

Given $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, pair (ij) 's bargaining problem is

$$\max_{\mathbf{x}_{ij} \in B_{ij}(\mathbf{x}_{-(ij)})} N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; b_{ij})$$

where $N_{ij} : X \times (0, 1) \rightarrow \mathbb{R}$ is a function of $\mathbf{x}$, parameterized by b_{ij} , such that

$$N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; b_{ij}) = \begin{cases} \left[\Delta\pi_{D_j}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}) \right]^{b_{ij}} \cdot \left[\Delta\pi_{U_i}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}) \right]^{1-b_{ij}} & \text{if } \mathbf{x}_{ij} \in B_{ij}(\mathbf{x}_{-(ij)}) \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Assumption 2 presents sufficient conditions for the existence of a continuous bargaining solution. These conditions are expressed in terms of gains from trade functions. We choose this approach because, although in Section 8 we show how to verify our assumptions, researchers may find alternative ways of proving Assumption 2 holds.

Assumption 2. Fix $\mathbf{b} \equiv \{b_{ij}\}_{(ij) \in G} \in (0, 1)^{|G|}$. Let $\mathcal{O} \subset \mathbb{R}^K$. For all (ij)

1. Firms' gains from trade are twice continuously differentiable in $\mathbf{x} \in \mathcal{O}$.
2. For all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, $\text{int}_{ij}(B_{ij}(\mathbf{x}_{-(ij)}))$ is nonempty.
3. For all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, N_{ij} has a unique global maximizer in $\mathbf{x}'_{ij} \in \text{int}_{ij}(B_{ij}(\mathbf{x}_{-(ij)}))$.
4. For all $\mathbf{x}_{-(ij)} \in \mathbb{R}^{K-(ij)}$, if $\mathbf{x}_{ij} \notin X_{ij}$, then $\Delta\pi_f(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}) \leq 0$ for at least one firm.

Section 8 shows how to verify Assumption 2; here we summarize those results. Continuous differentiability comes from the smoothness of the cost and revenue functions. These are standard properties in IO, including the models in every paper in Table 1 that uses Nash-in-Nash.

Continuous differentiability is violated if inputs sold by two suppliers are perfect substitutes for a retailer who buys from them. In that case, suppliers effectively engage in Bertrand competition in the bargaining stage. We present an example in Section 9.3. This situation is of no concern for models used in empirical IO because discrete choice and CES demand systems rule out perfect substitution.

Assumption 2.2 means there exists a jointly profitable contract. In Section 8.3 we show that this assumption holds in models with discrete choice demand, at least for Lebesgue-almost every $\theta \in \Theta$. Assumption 2.2 is implied by Assumption 2.3. We present them separately to clarify the discussion.

Assumption 2.3 states that each bargaining pair always has one contract that globally maximizes their Nash product. Strict quasi-concavity of $N_{ij}(\cdot, \mathbf{x}_{-(ij)}; b_{ij})$ over $B_{ij}(\mathbf{x}_{-(ij)})$ ensures this, but in most models it is not feasible to verify that quasi-concavity holds. We use our gradient descent algorithm to get around this problem, by showing that repeated failure of our algorithm to converge implies a violation of Assumption 2.3. We also note that an equilibrium may exist even if Assumption 2.3 fails.

Finally, Assumption 2.4 means inadmissible contracts are never jointly profitable. This ensures no Nash-in-Nash equilibria are ruled out by the construction of X . In Section 8.1 we show how to construct such admissible contracts sets.

Lemma 1. *Let Assumption 2 hold. For all $(ij) \in G$, for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, the bargaining problem has a unique solution, $\mathbf{x}_{ij}^* = \mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; b_{ij})$. $\mathbf{x}_{ij}^*$ is continuous in $(\mathbf{x}_{-(ij)}, b_{ij}) \in X_{-(ij)} \times (0, 1)$, parameterized by b_{ij} , and maps into the interior of X_{ij} . It is Pareto efficient, independent of irrelevant alternatives, and invariant to equivalent payoff representations.*

Proof. This and other proofs are relegated to Appendix A. The key, nontrivial step of the proof is showing lower hemicontinuity of the bargaining set correspondence.

Pareto efficiency is efficiency within the set of admissible contracts. Independence of irrelevant alternatives means that if, given $\mathbf{x}_{-(ij)}$ a pair bargains over contracts in two bargaining sets, $B_{ij}(\mathbf{x}_{-(ij)})$ and $B' \subsetneq B_{ij}(\mathbf{x}_{-(ij)})$, and $\mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; b_{ij}) \in B'$, then the bargaining solution when the bargaining set is B' is also $\mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; b_{ij})$. Invariance to equivalent payoff representations is invariance to positive affine transformations of profits.

Twice-continuous differentiability of gains from trade is relevant for the uniqueness proofs. For existence of a continuous bargaining solution only continuity is required.

Corollary 1. *All the results of Lemma 1 are preserved if Assumption 2.1 is weakened to continuity of gains from trade in $\mathbf{x} \in X$.*

5 Definition and Existence of Nash-in-Nash Equilibria

In this section we define a Nash-in-Nash equilibrium and use Assumption 2 to prove its existence.

Definition 1 (Nash-in-Nash Equilibrium). *Fix $\mathbf{b} \in (0, 1)^{|G|}$. A vector of contract terms $\mathbf{x} \in X$ is a Nash-in-Nash equilibrium if, for all (ij) , $\mathbf{x}_{ij} = \mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; b_{ij})$.*

Assign to each bargaining pair a number $g \in \{1, \dots, |G|\}$. Function $F : X \times (0, 1)^{|G|}$ is a function of $\mathbf{x}$, parameterized by $\mathbf{b}$, that stacks the bargaining solutions of each pair in increasing order of g . A Nash-in-Nash equilibrium is a fixed point of F .

Proposition 1. *Let Assumption 2 hold. Then, a Nash-in-Nash equilibrium exists.*

Corollary 2. *The result of Proposition 1 is preserved if Assumption 2.1 is weakened to continuity of gains from trade in $\mathbf{x} \in X$.*

6 Local and Global Uniqueness of Nash-in-Nash Equilibria

The first part of this section adds a differentiability assumption under which Nash-in-Nash equilibria are finite and locally unique. The second part presents an additional assumption under which the equilibrium is globally unique, and conditions that ensure that assumption holds. Finally, we present a test for the presence of multiple equilibria.

6.1 Finiteness and Local Uniqueness

Assumption 3. Fix $\mathbf{b} \in (0, 1)^{|G|}$. Assume

1. F is continuously differentiable in $(\mathbf{x}, \mathbf{b})$.
2. For all (ij) , if $\mathbf{x}$ is an equilibrium then $\partial x_{ijk}^* / \partial b_{ij} \neq 0$ for all $k \in \{1, \dots, K_{ij}\}$.

In Section 8 we show that continuous differentiability of F , like twice continuous differentiability of gains from trade, comes from the smoothness of cost and revenue functions, and mild regularity conditions on profits. Thus, a conclusion of Section 8 is that if gains from trade are twice-continuously differentiable, continuous differentiability of F is a weak additional assumption.

Assumption 3.2 states that in equilibrium all bargaining pair's contract terms depend nontrivially on their bargaining parameters. This statement, while intuitive, may require knowledge of the model at hand to be proven for all $\mathbf{x}$. Assumption 3.2, however, is not necessary for the proof, and merely rules out edge cases. In Section 8 we show that for models with discrete choice demand and constant returns to scale Assumption 3.1 holds and Assumption 3.2 is unnecessary.

Proposition 2. Let Assumptions 2 and 3 hold. Then for Lebesgue-almost every $\mathbf{b} \in (0, 1)^{|G|}$ there exists a finite number of locally unique Nash-in-Nash equilibria.

Proposition 2 shows that if bargaining solutions are continuously differentiable, then local uniqueness is a generic result, in the sense that it can only fail in a subset of Lebesgue measure zero of the parameter space. Therefore, the existence of infinite equilibria is an edge case result, which disappears under almost every arbitrarily small perturbation of $\mathbf{b}$.

6.2 Global Uniqueness

Notation. Let $D_{\mathbf{z}}h$ denote the array of first partial derivatives of a function h with respect to vector $\mathbf{z}$. Conditional on Assumptions 2 and 3.1 holding, Assumption 4 gives a necessary and sufficient condition for global uniqueness of a Nash-in-Nash equilibrium.

Assumption 4. If $\mathbf{x}$ is a Nash-in-Nash equilibrium, then $\det(I - D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b})) > 0$.

Proposition 3. Let Assumptions 2, 3.1, and 4 hold. Then there exists a globally unique Nash-in-Nash equilibrium.

Assumption 4 is not a condition on model primitives. Lemma 2 presents conditions under which Assumption 4 holds.

Lemma 2. Let Assumptions 2 and 3.1 hold. If

$$\sum_{(i'j') \neq (ij)} \sum_{k' \in \mathcal{K}_{(i'j')}} \left| \frac{\partial x_{i'j'k'}^*}{\partial x_{ijk}} \right| < 1 \quad (3)$$

holds for every (ij) , or

$$\sum_{(i'j') \neq (ij)} \sum_{k' \in \mathcal{K}_{(i'j')}} \left| \frac{\partial x_{ijk}^*}{\partial x_{i'j'k'}} \right| < 1 \quad (4)$$

holds for every (ij) , for every $k \in \mathcal{K}_{ij}$, then Assumption 4 holds.

The conditions of Lemma 2 have an economic interpretation: stability (Moulin, 1984; Escrihuela-Villar et al., 2024). That is, if cross effects between contracts are sufficiently small in every equilibrium, then the equilibrium is unique. Section 9.1.2 presents models with primitives such that the conditions of Lemma 2 are satisfied.

Stability is not necessary for global uniqueness. Lemma A.1 presents other sufficient, and necessary and sufficient, conditions for Assumption 4 to hold. These are conditions on the spectrum of F , so we relegate them to Appendix A.

In the introduction we stated that our results nest efficient contracts as a special case. Corollary 3 shows that if contracts are efficient, then the Nash-in-Nash equilibrium is globally unique.

Corollary 3. *Let Assumption 2 hold. If contracts are efficient, there is a globally unique Nash-in-Nash equilibrium.*

6.3 A Test for the Presence of Multiple Equilibria

Assumption 4 provides a test for the presence of multiple equilibria. The test is partial because it can reject global uniqueness, but cannot support it. Assuming a researcher has a sample that lets them observe an equilibrium $\mathbf{x}$, and a consistent estimate of F , Corollary 4 presents this test.

Corollary 4. *Let Assumptions 2 and 3.1 hold. If, at the observed equilibrium $\mathbf{x}$, $\det(I - D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b})) \leq 0$, then there are multiple equilibria.*

7 Computation Using Gradient Descent

In this section we show how to compute Nash-in-Nash equilibria using gradient descent, a derivative-based minimization algorithm. Gradient descent is a good fit for three reasons. First, under standard smoothness and step-size conditions used in modern optimization, gradient descent has a well developed convergence theory for non-convex problems: it always converges and, under random initialization, it almost surely converges to a local minimum and not a strict saddle point (Lee et al., 2016). These conditions are considerably weaker than the contraction type conditions that would guarantee convergence of fixed point iteration.

Second, in our setting Assumptions 2 and 3.1 imply that the objective function we construct below is continuously differentiable, so its gradient is well-defined and can be computed. Third, gradient-based methods scale well and often perform well in practice even when theoretical guarantees do not exist (Mnih and Salakhutdinov, 2007; Koren et al., 2009; Gemulla et al., 2011).

This section starts by explaining the gradient-based algorithm. Then we show that repeated evaluations of the algorithm using random starting points can speak to the uniqueness or multiplicity of equilibria. Finally, we compare gradient descent with alternative algorithms.

7.1 Algorithm Description

Given $\mathbf{b}$, consider the operator $\tilde{F} : X \rightarrow X$, defined in blocks for each (ij) as

$$\tilde{F}_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}) = \arg \max_{\hat{\mathbf{x}}_{ij} \in X_{ij}} N_{ij}(\hat{\mathbf{x}}_{ij}, \mathbf{x}_{-(ij)}; b_{ij})$$

That is, for each bargaining pair, $\tilde{F}_{ij}$ solves the bargaining problem without the participation constraints. Under Assumption 2.2 this relaxation is without loss. Consider the following function:

$$M(\mathbf{x}) = \frac{1}{2} \left\| \tilde{F}(\mathbf{x}) - \mathbf{x} \right\|_2^2.$$

M is non-negative and $M(\mathbf{x}) = 0$ if, and only if, $\mathbf{x}$ is an equilibrium. Hence, equilibrium computation reduces to minimizing M . Convergence to local minimizers that are not equilibria is immediately detected, so the strong global and local convergence results of gradient descent mean our algorithm is reliable. Since under Assumption 3.1 $\tilde{F}$ is continuously differentiable, the gradient of M is well-defined.⁶

By Lemma 1, $\tilde{F} = F(\cdot; \mathbf{b})$ and $D_{\mathbf{x}}\tilde{F} = D_{\mathbf{x}}F(\cdot; \mathbf{b})$. Therefore, by the chain rule the gradient of M is

$$D_{\mathbf{x}}M(\mathbf{x}) = (D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b}) - I)^T (F(\mathbf{x}; \mathbf{b}) - \mathbf{x}).$$

Gradient descent updates a candidate equilibrium using the gradient. Given a starting point $\mathbf{x}_0$ and learning rate $\alpha > 0$, we update the t -th element of the sequence by

$$\mathbf{x}_{t+1} = \mathbf{x}_t - \alpha D_{\mathbf{x}}M(\mathbf{x}_t) \tag{5}$$

Computing an equilibrium then reduces to computing M and its gradient. First, given $\mathbf{x}$, compute $M(\mathbf{x})$ by maximizing each $N_{ij}(\cdot, \mathbf{x}_{-(ij)}; b_{ij})$ using grid search to find $\tilde{F}_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)})$. The number of steps in this grid-search grows exponentially with $\max_{(ij) \in G} K_{ij}$. Thus, the discretization is confined to the inner bargaining problems, rather than to the full space X . Then, compute $D_{\mathbf{x}}M$ using automatic differentiation. Appendix B presents the technical details.

7.2 Searching for Multiple Equilibria

A common way of searching for multiple equilibria used in IO is running an algorithm many times using random, i.i.d. starting points $\mathbf{x}^0$. Repeated convergence to the same equilibrium is then heuristically interpreted as increasing one's confidence in the global uniqueness of the equilibrium.

Nehzati and Cussen (2026) present conditions under which repeated convergence to the same outcome can be *formally* interpreted, using Bayesian inference, as evidence for the global uniqueness of an equilibrium. The relevant conditions are measurability of the function that maps initial conditions into an equilibrium, finiteness of the set of equilibria, and positive Lebesgue measure of each equilibrium's

⁶Convergence results for gradient descent require twice continuously differentiable objective functions, meaning Assumption 3.1 must be extended to twice continuous differentiability. See Section 8.5 for the proof of twice continuous differentiability of F with respect to $\mathbf{x}$.

basin of attraction.⁷ These sufficient conditions hold for our algorithm: Nehzati and Cussen (2026) show gradient descent is measurable, finiteness comes from Proposition 2, and positive Lebesgue measure of the basins of attraction comes from the local convergence property of gradient descent. Therefore, repeated i.i.d. restarts reveal the distribution of computed equilibria.

Moreover, in our setting every reachable equilibrium with positive-measure basin of attraction is found almost surely under infinitely many restarts. If every equilibrium is reachable by the algorithm, repeated convergence to the same equilibrium has a disciplined interpretation as evidence in favor of global uniqueness.

7.3 Alternative Algorithms

7.3.1 Fixed Point Iteration

In a general vertical relations setting, fixed point iteration is not guaranteed to be a contraction—otherwise the proofs of existence and uniqueness of equilibrium would be trivial. This is where gradient descent has an advantage: the conditions under which it is well behaved are broader than the conditions that would guarantee convergence of fixed point iteration. The lack of convergence results for fixed point iteration, however, is compensated by its lower per-step computational cost, as fixed point iteration does not require computing gradients.

In Section 9.2 we compare the speed of gradient descent and fixed point iteration. In Grennan (2013)’s model of the market for stents fixed point iteration converges, and quickly.⁸ However, the gains in computational speed seem small enough that gradient descent, with its convergence properties, is still an attractive alternative.

7.3.2 Homotopy Methods

Homotopy methods use the solution of an easy case of a complex system of equations to get the solution of a harder case by moving through the parameter space. These methods have proven successful in dynamic games (Borkovsky et al., 2010). For the vertical relations games we consider, a natural starting point is some $\mathbf{b} \in \{0, 1\}^{|G|}$ (a model with take-it-or-leave-it offers).

Homotopy methods, however, can systematically fail to compute certain equilibria. That is, there may exist an equilibrium vector $\mathbf{x}^*$ such that no starting point $\mathbf{x}_0 \neq \mathbf{x}^*$ will lead to $\mathbf{x}^*$ being the computed equilibrium using the homotopy method. By contrast, Section 7.2 shows that repeated restarts of gradient descent admit a clear interpretation for all reachable equilibria with positive-measure basins of attraction. Moreover, homotopy methods are slow compared to gradient descent. All of this means that, in our vertical relations setting, homotopy methods are not as convenient as gradient descent, especially when searching for multiple equilibria.

⁷An equilibrium $\mathbf{x}^*$ ’s basin of attraction under computation algorithm Ψ , $ba^\Psi(\mathbf{x}^*)$, is the subset of X such that, if Ψ is initialized at an point $\mathbf{x}_0 \in ba^\Psi(\mathbf{x}^*)$, the output of the algorithm is $\mathbf{x}^*$.

⁸Results available upon request.

7.3.3 Full Grid Search

Full grid search works as follows: for each (ij) , for each $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, evaluate $N_{ij}(\cdot, \mathbf{x}_{-(ij)}; b_{ij})$ at each $\mathbf{x}_{ij} \in B_{ij}(\mathbf{x}_{-(ij)})$. Use these computations to construct the bargaining solution for each (ij) , and find an equilibrium. This can be done using fixed point iteration.

Full grid search’s main disadvantage is its slowness. Given a number of gridpoints for each dimension of X , the search cost grows exponentially with K , because every possible combination of contract terms must be explored for every bargaining pair. This is in sharp contrast with gradient descent, which only evaluates Nash products along the path of updates. This means that gradient descent’s search cost grows exponentially with $\max_{(ij) \in G} K_{ij}$ —considerably slower.

Unlike other methods, full grid search gives researchers a numerical representation of the functional forms of the Nash products and the bargaining solutions. Thus, without providing a proof, the output of a full grid search gives researchers numerical information about a model—particularly about Assumption 2.3.

To illustrate this advantage of full grid search, Appendix C presents a model with two multiproduct retailers who negotiate linear input prices with two suppliers each. To increase the complexity of the model, the demand system is mixed logit. With the caveat that this is just a numerical exercise, two results emerge. First, fixed point iteration starting from *any* point in X always yields the same computed equilibrium. Second, the Nash product functions not only have a unique global maximizer, but are also hump-shaped. Hump-shaped Nash products are expected if contracts are efficient, but there is nothing to suggest that result in highly nonlinear problems. We obtain these results for different parameter values and grids of X .

8 Verifying Assumptions

This section shows how to verify Assumptions 2 and 3 using model primitives, or at least test for failures in the case of Assumption 2.3. We begin with the construction of an admissible contracts set X , and then follow the order in which the assumptions were introduced. The section ends with a checklist empirical researchers can go through to verify all the sufficient assumptions for existence and uniqueness hold in their models. The results of this section are particularly focused on models with discrete choice demand and constant returns to scale. Researchers who prefer to skip the technical parts of the section can skip to that last subsection.

8.1 Construction of the Admissible Contracts Set

This subsection shows how to construct a nonempty, compact, and convex admissible contract set X that satisfies Assumption 2.4 (no jointly profitable contracts are excluded). The results are contingent on contracts being of one of the four most common types observed in the empirical literature: linear input prices, two-part tariffs, revenue sharing, and quantity-forcing, but may be extended to other types of contract.

The only primitive our proof requires is that agents in the model face a resource constraint. This assumption is natural, and indeed unavoidable in applied work (Morrow, 2015). What follows are

examples of how to construct X under each contract type.

1. Linear input prices. A contract $\mathbf{x}_{ij} \in X_{ij} \subset \mathbb{R}^{K_{ij}}$ consists of linear input prices if each element is a constant unit price. For such a contract to be profitable, each price of a unit of input, x_{ijk} , must be less than (a large multiple of) total industry revenue, which is finite. So, industry revenues give an upper bound for each x_{ijk} . The negative of the upper bound provides a lower bound.⁹

2. Two-part Tariffs. A two-part tariff is a contract with $K_{ij} - 1$ (constant) unit prices and a lump-sum transfer. The bounds on unit prices used for linear input prices apply to both the linear prices and the lump-sum transfer of a two-part tariff.

3. Revenue Sharing. Revenue sharing involves transferring to the supplier a fraction of the revenue obtained from the sale of a product. Even if subsidies (i.e., revenue shares below 0% or above 100%) are allowed, finite resources means the shares must be bounded.

4. Quantity-forcing. A quantity-forcing contract consists of $K_{ij} - 1$ specified quantities for $K_{ij} - 1$ inputs plus and a lump-sum transfer. Large multiples of maximum industry revenue bound the transfers from above and below. Zero bounds quantities from below, and the upper bound is imposed by the upper bound for the transfer and a positive marginal cost of production for the retailer (or an upper bound on quantity demanded, as in discrete choice models).

Lemma 3. *Let resources be finite. If contracts are linear input prices, two-part tariffs, revenue sharing, or quantity-forcing, then there exists a nonempty admissible contracts set that is compact, convex, and satisfies Assumption 2.4.*

Proof. See the paragraphs above. ■

8.2 Twice Continuous Differentiability of Gains From Trade

It suffices to prove that on-path profits are twice continuously differentiable in $\mathbf{x}$ and that off-path profits are twice-continuously differentiable in $\mathbf{x}_{-(ij)}$. To set up the proof, note that Assumption 1 ensures that, for each $\mathbf{x} \in X$, in an interior second stage equilibrium each retailer's first-order condition with respect to $\mathbf{p}_j$ holds:

$$D_{\mathbf{p}_j} \pi_{D_j}(\mathbf{p}_j^*, \mathbf{p}_{-j}^*, \mathbf{x}) = 0 \quad \forall j \in \mathcal{J}. \quad (6)$$

We collect the first-order conditions of all retailers into one system of equations, $[D_{\mathbf{p}_j} \pi_{D_j}(\mathbf{p}_j, \mathbf{p}_{-j})]_j = 0$. We call this the second stage equilibrium system.

Assumption 5. *Assume that*

1. *All firms' profit functions, $\pi_f(\mathbf{p}, \mathbf{x})$, are four times continuously differentiable in $(\mathbf{p}, \mathbf{x})$.*
2. *At the second stage equilibrium, $\mathbf{p}^*$ (6) holds and its Jacobian with respect to $\mathbf{p}$ is invertible:*

$$\det \left(D_{\mathbf{p}} [D_{\mathbf{p}_j} \pi_{D_j}(\mathbf{p}_j^*, \mathbf{p}_{-j}^*)]_j \right) \neq 0. \quad (7)$$

⁹Under CES demand expenditure in an industry is exogenous, making the proof of Lemma 3 below trivial.

Four times continuous differentiability with respect to the strategic variable holds in standard IO models, including CES, logit, nested logit, and mixed logit demand systems with linear-in-price utility. See Konovalov and Sándor (2010), Morrow and Skerlos (2010b), Cussen (2026), and Appendix E.

Assumption 5.2 is a standard condition used to prove uniqueness of interior equilibria (Konovalov and Sándor, 2010). Given Assumption 1, failure of (7) means the second stage equilibrium system is degenerate at the equilibrium, yet the equilibrium is still unique. Assumption 5.2, which is sufficient but not necessary for our proof, rules out such pathological cases, and is known to hold in models with CES, logit, and mixed logit with linear utility (Konovalov and Sándor, 2010; Cussen, 2026).

Proposition 4. *Let Assumptions 1 and 5 hold. Then, for all $(ij) \in G$, all firms' gains from trade functions are three times continuously differentiable with respect to $\mathbf{x} \in \mathcal{O} \supset X$.*

We want three times continuous differentiability of gains from trade to get twice continuous differentiability of bargaining solutions. Let us use the strongest convergence results for gradient descent. The proofs of local and global uniqueness, however, only require the twice continuous differentiability of Assumption 2.1.

Corollary 5. *Let Assumptions 1 and 5.2 hold, and relax Assumption 5.1 to three times continuous differentiability. Then, for all $(ij) \in G$, all firms' gains from trade functions are twice continuously differentiable with respect to $\mathbf{x} \in \mathcal{O} \supset X$.*

8.3 Nonempty Interior of the Bargaining Set

This subsection has two parts. First, we give conditions under which the bargaining set is always nonempty. We then extend the proof to nonemptiness of the interior of the bargaining set.

8.3.1 Nonempty Bargaining Set

We prove nonemptiness of the bargaining set for discrete choice models. Let contracts consist of linear input prices, two-part tariffs, or revenue sharing. We skip the trivial proof for quantity-forcing contracts: all that is required is setting quantities and the lump-sum transfer to zero.

Let the demand system be logit, nested logit, or mixed logit, and let retailers choose prices. Following the empirical literature, we assume firms have constant returns to scale and that their cost functions are additively separable by product. Finally, we assume consumers face a budget constraint and that equilibrium markups are non-negative. Proofs of non-negativity of markups can be found in Morrow and Skerlos (2010b,a), Morrow (2015), and Liu (2006), among other papers.

Lemma 4. *Consider the models described above. Let contracts be linear input prices, two-part tariffs, or revenue sharing, and let consumers have finite income. Let Assumptions 1 and 2.1 hold. Then, for all $(ij) \in G$, for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, there exists $\mathbf{x}$ that gives both firms nonnegative gains from trade.*

8.3.2 Nonempty Interior

To prove our nonempty interior result, we note that the bargaining set admits two plausible alternative, closely related formulations. Making the dependence of gains from trade on $\boldsymbol{\theta}$ explicit, these are

$$\tilde{B}_{ij}(\mathbf{x}_{-(ij)}) = \left\{ \mathbf{x}_{ij} \in X_{ij} : \Delta\pi_{D_j}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) \geq 0 \text{ and } \Delta\pi_{U_i}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) \geq 0 \right\} \quad (8)$$

and

$$\hat{B}_{ij}(\mathbf{x}_{-(ij)}) = cl_{ij} \left\{ \mathbf{x}_{ij} \in X_{ij} : \Delta\pi_{D_j}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) > 0 \text{ and } \Delta\pi_{U_i}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) > 0 \right\} \quad (9)$$

Expression (8) defines the bargaining set as the set of contracts that give both firms nonnegative gains from trade, while (9) is the closure of the set of jointly profitable contracts.

Ex ante, (8) and (9) may seem quite different from each other and from our definition of gains from trade. This difference could potentially affect our results, but that turns out not to be the case. In Appendix F (Theorem F.1) we show that under a proper parameterization of profit functions the three definitions of the bargaining set are identical for Lebesgue-almost every parameter vector $\boldsymbol{\theta} \in \Theta$. So, for almost every $\boldsymbol{\theta} \in \Theta$, if the bargaining set is nonempty its interior is nonempty too.

To apply this result we prove that models with logit and mixed logit demand and linear-in-price utility and constant returns to scale satisfy the proper parameterization conditions (Proposition F.1 and Corollary F.1).¹⁰ Given the technical nature of Appendix F, in this subsection we present our main assumption and explain how logit and mixed logit are properly parameterized.

Assumption 6. *For all $(ij) \in G$, for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, $\Delta\pi_{U_i}(\cdot, \mathbf{x}_{-(ij)}; \cdot)$ and $\Delta\pi_{D_j}(\cdot, \mathbf{x}_{-(ij)}; \cdot)$ are K_{ij} times continuously differentiable with respect to $(\mathbf{x}_{ij}, \boldsymbol{\theta})$.*

The smoothness of profit functions with respect to strategies, contracts, and parameters under discrete choice demand systems is shown in Appendix E.

Proper parameterization requires that perturbations of different parameters in $\boldsymbol{\theta}$ perturb gains from trade (and their gradients) in different directions. In logit and mixed logit with linear-in-price utility and constant returns to scale this requirement is satisfied by arbitrarily small perturbations of consumers' mean utilities and retailers' marginal cost functions.

8.4 Unique Global Maximizer of the Nash Product

Because most demand systems used in empirical IO (i.e., discrete choice) have no closed form, in general it is unfeasible to prove that the Nash product has a unique global maximizer (Assumption 2.3). For such cases, in this subsection we show how to use our gradient descent algorithm to test whether Assumption 2.3 holds.

Under Assumption 2.1, for all (ij) , for all $\mathbf{x}_{-(ij)}$, the bargaining solution correspondence is nonempty-valued.¹¹ This correspondence may have a fixed point even if it is not convex-valued.¹² Therefore,

¹⁰Due to space constraints we do not extend our results to nested logit. However, since logit is a special case of nested logit, a tedious re-writing of the proof of Corollary F.1 would extend our results to nested logit. In Section 9.2 we prove nonemptiness of the interior of the bargaining set with random coefficients nested logit demand directly.

¹¹Herrero (1989) provides a noncooperative foundation for bargaining problems where the Nash bargaining solution is not unique.

¹²By the extreme value theorem there is always at least one bargaining solution. By the maximum theorem, the bargaining solution correspondence is upper hemicontinuous (see the proof of Lemma 1 in Appendix A). The same lack of

existence, local uniqueness, and global uniqueness of Nash-in-Nash equilibria can survive violations of Assumption 2.3.

Before moving on, it is important to point out that Proposition 6 below shows that, for Lebesgue-almost every $\theta \in \Theta$, for models with discrete choice demand, constant returns to scale, and linear input prices, the critical points of the Nash product are always isolated. Therefore, if Assumption 2.3 fails, the bargaining solution will not be convex-valued, so there is no ex ante way to show it has a fixed point. More important, in general there will be no continuous selection from the elements of the correspondence.

With these properties of the bargaining solution correspondence in mind, the following paragraphs explain what we can learn about Assumption 2.3 by running our algorithm multiple times.

Case 1. Assumption 2.3 never fails. As explained in Section 7.2, in this case every equilibrium will eventually be computed using our algorithm.

Case 2. Assumption 2.3 fails but there is a Nash-in-Nash equilibrium. Gradient descent requires a continuously differentiable objective function. If Assumption 2.3 fails, gradient descent can only work, and converge to an equilibrium, if (1) a continuously differentiable bargaining solution function can be selected from the bargaining solution correspondence, and (2) gradient descent is selecting this function along its path of updates, defined by a starting point $\mathbf{x}_0$ and update rule (5). Since neither of these conditions is guaranteed, convergence of gradient descent when Assumption 2.3 fails is an edge case.¹³

As explained above, for models with discrete choice demand, constant returns to scale, and linear input prices Proposition 6 implies that in general there is no continuous selection. Therefore, convergence of our algorithm is particularly strong evidence for Assumption 2.3 in these models.

Case 3. Assumption 2.3 fails and there is no Nash-in-Nash equilibrium. In this case gradient descent will not work, as it is not possible to construct a continuously differentiable objective function. Attempts to use our gradient descent algorithm will never converge.

Conclusion. In all but edge cases, convergence of our algorithm implies that Assumption 2.2 is satisfied. In the edge cases convergence implies that a Nash-in-Nash equilibrium exists even if Assumption 2.2 fails, and search for multiple equilibria can be performed as explained in Section 7.2.

To show directly that uniqueness of the bargaining solution can hold in models with discrete choice demand, in Appendix G we prove that result for a model with two suppliers, two single product retailers, and logit demand.

closed form that precludes a proof of uniqueness of the global maximizer of the Nash product, however, precludes a proof that the bargaining solution correspondence is convex-valued, so we cannot use the Kakutani fixed point theorem to prove existence of a Nash-in-Nash equilibrium.

¹³See for example the Michael and Deutsch-Kenderov selection theorems. The conditions of those theorems may be unverifiable or may fail to hold. Moreover, the theorems guarantee existence of a continuous selection function, but say nothing about continuous differentiability.

8.5 Continuous Differentiability of the Bargaining Solution

Assumption 7. For all $(ij) \in G$, for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, the Hessian of $N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; b_{ij})$ is invertible at the bargaining solution $\mathbf{x}_{ij}^* = \mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; b_{ij})$:

$$\det \left(D_{\mathbf{x}_{ij} \mathbf{x}_{ij}}^2 N_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; b_{ij}) \right) \neq 0.$$

Assumption 7 is weak for two reasons. First, Morse's Lemma implies that functions whose Hessians are non-invertible at critical points are pathological cases that disappear under arbitrarily small perturbations. In settings with gains from trade continuously differentiable in $(\mathbf{x}, \boldsymbol{\theta})$, that may be achieved with arbitrarily small perturbations of $\boldsymbol{\theta}$.¹⁴ Proposition 6 below shows how such an argument from perturbations can be used to obtain invertibility of the Hessian for a large class of models.

Second, Assumption 7 is not necessary: a function's maximizer can be infinitely continuously differentiable even if the sufficient second order condition does not hold.¹⁵

Proposition 5. Let Assumptions 2 and 7 hold. Then, for all $(ij) \in G$, $\mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; b_{ij})$ is continuously differentiable with respect to $(\mathbf{x}_{-(ij)}, b_{ij})$. If we strengthen Assumption 2.1 to three times continuous differentiability (Assumptions 1 and 5 and Proposition 4), $\mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; b_{ij})$ is twice continuously differentiable with respect to $(\mathbf{x}_{-(ij)}, b_{ij})$.

The following proposition proves continuous differentiability without imposing Assumption 7. Let the demand system be logit, nested logit, mixed logit or RCNL, and let retailers compete in prices. All firms have constant returns to scale, and all contracts are linear input prices. Let $\mathbf{c}$ be the vector of (nonnegative) supplier marginal cost of production of all inputs. Without loss, we assume each marginal cost is bounded from above by some $\bar{c} > 0$, which may be arbitrarily high.

Proposition 6. Let Assumption 2 hold and $\mathbf{b} \in (0, 1)^{|G|}$. Then, for Lebesgue-almost every $\mathbf{c} \in [0, \bar{c}]^K$ Assumption 7 holds and $F(\mathbf{x}; \mathbf{b})$ is continuously differentiable with respect to $(\mathbf{x}_{-(ij)}, b_{ij})$. If we strengthen Assumption 2.1 to three times continuous differentiability (Assumptions 1 and 5 and Proposition 4), $F(\mathbf{x}_{-(ij)}; \mathbf{b})$ is twice continuously differentiable with respect to $(\mathbf{x}_{-(ij)}, \mathbf{b})$.

Proposition 6 applies to every paper in Table 1 that uses a discrete choice demand system, making it a powerful result for empirical research.

8.6 Dependence of Equilibrium Contracts on Bargaining Power

Assumption 3.2 states that, in a Nash-in-Nash equilibrium, for each pair (ij) , all its contract terms depend nontrivially on (ij) : $\partial x_{ijk}^* / \partial b_{ij} \neq 0$ for all $k = 1, \dots, K_{ij}$. Lemma 5 presents primitives under which Assumption 3.2 holds. These include models with no closed form solution.

Lemma 5. Let Assumptions 2 and 3.1 hold and $K = |G|$. For each $(ij) \in G$, for each $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, let $\Delta \pi_f(\cdot; \mathbf{x}_{-(ij)})$ be strictly monotonic in x_{ij} for some $f \in \{U_i, D_j\}$. Then, Assumption 3.2 holds.

¹⁴See chapter 1 of Guillemin and Pollack (1974) for an introduction to Morse theory and a proof of Morse's Lemma.

¹⁵Let $\boldsymbol{\theta} \in \mathbb{R}^K$ and $f(\mathbf{x}; \boldsymbol{\theta}) = \sum_{k=1}^K -(x_k - \theta_k)^4$. Its argmax is $x_k(\boldsymbol{\theta}) = \theta_k$ for all k , infinitely continuously differentiable in $\boldsymbol{\theta}$. But all the second partial derivatives are 0 at the argmax, so the sufficient second-order condition never holds.

To state the conditions of Lemma 5 in terms of primitives, note that if contracts are linear input prices, two-part tariffs, or revenue sharing, then D_j 's gains from trade are strictly decreasing in every element of $\mathbf{x}_{ij}$. $K_{ij} = 1$ for all (ij) is a common assumption in empirical work.¹⁶

The extension of Lemma 5 to $K > |G|$, while intuitive, requires conditions on the curvature of $N_{ij}(\cdot, \mathbf{x}_{-(ij)}; b_{ij})$ to rule out edge cases. Thus, even though violations of Assumption 3.2 seem pathological, Lemma 6 presents alternative, general conditions under which local uniqueness of Nash-in-Nash equilibria holds in models with discrete choice demand.

Lemma 6. *Consider the models treated in Proposition 6. Let $\mathbf{c} \in \mathbb{R}^K$ be a vector of suppliers' marginal costs. Let Assumptions 2 and 3.1 hold. Then, for each $\mathbf{c}$ there exists an open neighborhood that contains it, $\mathcal{C}(\mathbf{c})$, such that for Lebesgue-almost every $\mathbf{c}' \in \mathcal{C}(\mathbf{c})$ the model with supplier marginal cost vector $\mathbf{c}'$ has a finite number of locally unique Nash-in-Nash equilibria.*

8.7 A Checklist for Researchers

Empirical researchers wishing to verify our sufficient conditions for existence and (at least local) uniqueness of Nash-in-Nash equilibria can be satisfied by answering the following questions:

1. Does the model have a discrete choice demand system with linear utility, constant returns to scale, and linear input prices?
2. Does a unique pure strategy equilibrium exist in the second stage?
3. Are contracts linear input prices, two-part tariffs, or revenue sharing?
4. Are profit functions (at least) $\max\{4, K_{ij}\}$ times continuously differentiable in strategies, contract terms, and parameters $\boldsymbol{\theta}$ for all $(ij) \in G$?
5. How many inputs do firms in each bargaining pair trade? Are the gains from trade of one firm in each pair strictly monotonic in each element of the contract?
6. Are there concerns that violations of Assumption 2.3 may lead to nonexistence?

Answers. If the answer to question 1 is yes, then the results of this section tell us that the only Assumptions we cannot guarantee ex ante, at least for Lebesgue-almost every parameter vector, are Assumption 2.3, Assumption 4 and, in some models with random coefficients, Assumption 1. This leads us to question 2, which is treated in Section 3.2. That section presents a list of existence and uniqueness results researchers may use, including for complicated cases of mixed logit demand.

If the answer to question 3 (contract type) is yes, then constructing an admissible contracts set X that satisfies Assumption 2.4 is easy: see Lemma 3. If the contracts are of some other type the proof of Lemma 3 provides guidance on how to construct X .

Even if the answer to question 1 is no, all standard demand systems answer yes to question 4 (continuous differentiability).

If the answers to both parts of question 5 are yes, then Lemma 5 ensures that bargaining solutions depend nontrivially on the bargaining parameter. If one of the answers is no, then that assumption

¹⁶See Grennan (2013, 2014), Ho and Lee (2017, 2019), Noton and Elberg (2018), Crawford et al. (2018), and Alvarez et al. (2023).

may fail in edge cases. For the models described in question 1 this is not relevant, as Lemma 6 provides an alternative way to obtain local uniqueness.

Finally, for models without closed form, Section 8.4 shows how to use our algorithm to test for failures of uniqueness of the bargaining solution that lead to nonexistence of Nash-in-Nash equilibria (question 6).

9 Examples

We present three examples. First, we cover the most common specification in Table 1: mixed logit models with linear-in-price utility, constant returns to scale, and linear input prices. Using our results in this paper, we go through the checklist in Section 8.7. We also provide a family of primitives under which a globally unique Nash-in-Nash equilibrium exists.

Second, we present Grennan (2013)’s model of the market for stents, and use it to show the performance of our algorithm. Finally, we present a modification of Horn and Wolinsky (1988)’s model with discontinuities in firms’ gains from trade. This example illustrates why discontinuities are not a concern in standard models used in empirical IO, and why the number of bargaining pairs ($|G|$) and contract terms (K) in the models we consider are generally flexible.

9.1 Mixed Logit

Consider mixed logit models with multiproduct firms, constant returns to scale, and linear utility. Every paper in Table 1 that uses mixed logit has this specification. Here we show that, if contracts are linear input prices, these models can satisfy our conditions for local uniqueness. Extension to two-part tariffs and other demand systems (logit, nested logit, RCNL, CES) is straightforward. Without loss, we only look at on-path profits. The analysis for off-path profits is the same.

9.1.1 General Case

Each retailer sells a set of products L_j , $|L_j| \geq 1$. Every product is different, and each requires at least one input for its production. Consumers can buy an outside good. Retailers compete by setting prices independently. We set the market size to one and ignore fixed costs, which drop out of gains from trade (Lemma 1). Letting s_l be the market share of good $l \in L_j$, retailer profits are

$$\pi_{D_j}(\mathbf{p}_j, \mathbf{p}_{-j}; \mathbf{x}, \boldsymbol{\theta}) = \sum_{l \in L_j} (p_l - mc_l(\mathbf{x}, \boldsymbol{\theta})) s_l(\mathbf{p}, \boldsymbol{\theta}). \quad (10)$$

where mc_l is the marginal cost of good l as a function of contracts. The dependence of market shares and marginal costs on $\boldsymbol{\theta}$ is explained after presenting suppliers’ profits. Following the literature, marginal costs satisfy

$$\iota(mc_l) = \gamma_l^0 + \mathbf{x}_{jl}' \boldsymbol{\gamma}_l^1 + \nu_l' \boldsymbol{\gamma}_l^2 + \omega_l \quad (11)$$

where $\boldsymbol{\gamma} = (\gamma_l^0, \gamma_l^1, \gamma_l^2)$ is a product-specific parameter vector, $\mathbf{x}_{jl}$ is the vector of unit costs for the inputs D_j needs to sell l , and ν_l and ω_l are observed and unobserved shocks. The functions ι commonly

employed in the literature are the identity and the log functions (Conlon and Gortmaker, 2020), so we restrict our attention to those.

Moving on to suppliers' profits, let H_i be the set of inputs U_i supplies. Let the sets $\{H_i\}_{i \in \mathcal{I}}$ form a partition of the input space. Because retailers have constant returns to scale, the amount of input $h \in H_i$ that retailer D_j buys to produce good $l \in L_j$ is linear in s_l , with transformation factor $a_{jlh} \geq 0$ determined by the parameters in (11). So, suppliers' profits in the second stage are

$$\pi_{U_i}(\mathbf{p}, \mathbf{x}) = \sum_{j: (ij) \in G} \sum_{l \in L_j} \sum_{h \in H_i} a_{jlh} s_l(\mathbf{p}; \boldsymbol{\theta}) (x_h - c_h(\boldsymbol{\theta})). \quad (12)$$

where $c_h \geq 0$ is U_i 's marginal cost of producing input h . Given these functional forms, $\boldsymbol{\theta}$ collects parameters in consumers' utility functions and all firms' marginal costs.

Now we answer the questions in section 8.7. Regarding existence and uniqueness of Bertrand-Nash equilibrium in the second stage, see Aksoy-Pierson et al. (2013) and Cussen (2026), and Section 3.2 for a discussion. As for question 3, we are assuming contracts are linear input prices.

Appendix E uses existing results to show that profit functions with mixed logit demand and constant returns to scale satisfy the continuous differentiability conditions of Assumptions 5.1 and 6.

If all bargaining pairs trade one input, then Lemma 5 proves that bargaining solutions depend nontrivially on b_{ij} . For $K_{ij} > 1$ for some pair, Lemma 6 yields a local uniqueness result without Assumption 3.2.

Violations of Assumption 2.3 are detected by our gradient descent algorithm (question 5), as shown in Section 9.2. In Section 9.1.2 we present a subset of the class of models considered here for which Assumption 4 is known to hold, guaranteeing a globally unique Nash-in-Nash equilibrium. Table 2 summarizes the results of this subsection

9.1.2 Global Uniqueness in Mixed Logit Models

Now we present a non-trivial example with mixed logit demand and a globally unique Nash-in-Nash equilibrium. Following the previous subsection, we let Assumptions 2 and 3.1 hold. We add the following restriction: for each bargaining pair $K_{ij} = 1$ and U_i 's gains from trade are

$$\Delta \pi_{U_i}(\mathbf{x}) = \sum_{l \in L_j} a_{jlh} s_l(\mathbf{x}) (x_h - c_h). \quad (13)$$

At least two kinds of models satisfy (13). First, models where each supplier sells to only one retailer. Second, "gatekeeper" models, where retailers act as buyers for final consumers (Gowrisankaran et al., 2015).¹⁷

Lemma 7. *Consider the model described above. Let Assumptions 2 and 3.1 hold. Then, there exists a subset of $(0, 1]^{|G|}$ of positive Lebesgue measure such that, for all $\mathbf{b}$ in that subset, Assumption 4 holds and the Nash-in-Nash equilibrium is globally unique.*

¹⁷Gatekeeper models are particularly well suited to the US healthcare industry, and may extend to other settings (Gowrisankaran et al., 2015).

The proof of Lemma 7 does not depend on the demand system being mixed logit, so it extends to other discrete choice demand systems.

Lemma 7 suggests a wider condition under which the Nash-in-Nash equilibrium is globally unique. As the diversion ratios from any good l to the outside good grow relative to diversion ratios to all other inside goods, for all bargaining pairs suppliers' gains from trade will converge to (13) (because upon disagreement the sales of other goods for which U_i sells inputs will be relatively unaffected).¹⁸ We therefore conjecture that if diversion ratios to the outside good are sufficiently high the conditions of Lemma 2 will hold.

Table 2: Verifying Assumptions in Mixed Logit Models

Assumption	
Assumption 1	Aksoy-Pierson et al. (2013), Cussen (2026)
Assumption 2.1	Assumptions 1 and 5, and Proposition 4
Assumption 2.2	Lemma 4, Corollary F.1
Assumption 2.3	See Section 9.2
Assumption 2.4	$X = [-W, W]^K$, W = total wealth (Lemma 3)
Assumption 3.1	Assumptions 2 and 7 + Proposition 5, or Proposition 6
Assumption 3.2	Lemma 5, Lemma 6
Assumption 4	See Sections 9.1.2 and 9.2
Assumption 5.1	Cussen (2026), Appendix E
Assumption 5.2	Cussen (2026)
Assumption 6	Appendix E
Assumption 7	Proposition 6

Notes: Uniqueness of the bargaining solution (Assumption 2.3) is not verified analytically for the general case; in general, our algorithm is required. See Section 9.2 for Assumption 2.3 and section 9.1.2 for models that satisfy Assumption 4.

9.2 Applying our Algorithm: The Market for Stents

Grennan (2013) presents a model of the wholesale market for stents. The formal treatment is in Appendix D.

Each market has a single hospital (retailer D_j) that buys stents from multiple suppliers. Doctors in a hospital choose whether to perform an angioplasty on a patient and, if so, which stent to use. Hospitals negotiate linear stent prices with producers, and receive a reimbursement for each angioplasty performed. Each stent price is negotiated separately, so $K_{ij} = 1$ for all pairs.

Patients and doctors' preferences are aligned and represented by doctors' utility functions, which are linear and decreasing in the cost of the stent for the hospital. This is a discrete choice problem so, conditional on stent prices, doctors' choices are unique. This makes the model a "gatekeeper" model,

¹⁸The diversion ratio between products l and $k \in L \setminus \{l\}$ is the fraction of consumers who stop consuming l after an increase in p_l and switch instead to k .

and reduces the second stage to choosing between finite alternatives.

The marginal cost of producing a stent i is c_i . Let $X_{ij} = [c_i, r_i]$ for all (ij) , where r_i is the greatest reimbursement the hospital can get for stent i . This admissible contract set guarantees no Nash-in-Nash equilibria are excluded.

Let the vector of parameters θ collect suppliers' marginal costs. With RCNL discrete choice demand and constant returns to scale, all demand and cost functions are infinitely continuously differentiable with respect to $(\mathbf{x}, \theta)$. Lemma 8 states that, if reimbursements can be sufficiently high, jointly profitable contracts always exist. This proves directly that Assumption 2.2 holds in this model.

Lemma 8. *There exist vector $(\tilde{r}_i)_{i \in \mathcal{I}}$ such that, if $(r_i)_{i \in \mathcal{I}}$ is such that $r_i \geq \tilde{r}_i$ for all i , Assumption 2.2 holds in this model.*

So far, given the results for discrete choice models in Section 8, we have verified all but one (resp. two) of the sufficient assumptions for local (resp. global) uniqueness; Assumptions 2.3 and 4 remain (see Table 3).

Table 3: Verifying Assumptions in Grennan 2013

Assumption	
Assumption 1	Given $\mathbf{x}$, doctor's choice is unique
Assumption 2.1	Assumptions 1 and 5, and Proposition 4
Assumption 2.2	Lemma 8
Assumption 2.3	Application of our algorithm
Assumption 2.4	$X_{ij} = [c_i, r_i]$ for all (ij)
Assumption 3.1	Assumptions 2 and 7 + Proposition 5, or Proposition 6
Assumption 3.2	Lemma 5
Assumption 5.1	Appendix E
Assumption 5.2	Trivial: second stage is discrete choice
Assumption 6	Appendix E
Assumption 7	Proposition 6

Notes: Assumption 4 cannot be verified directly; see Section 7.2 for how to use our algorithm.

We compute Nash-in-Nash equilibria, and test for violations of Assumption 2.3, using our gradient descent algorithm on a grid of X with 10^{24} points. To search for multiple equilibria we do the computation 100 times using randomly selected starting points. Panel A of Table 4 shows the results.

Several conclusions emerge. First, our algorithm is indeed fast. Despite the grid having 10^{24} points, the algorithm takes less than 20 seconds to converge. Second, it always converges to an equilibrium. Third, because the algorithm always converges to an equilibrium, any failure of Assumption 2.3 does not lead to nonexistence of equilibrium.

Fourth, because gradient descent eventually computes all equilibria, using the framework of Nehzati and Cussen (2026) our computations provide strong evidence in favor of the global uniqueness of the

Nash-in-Nash equilibrium. Panel B of Table 4 shows descriptive statistics of the distribution of computed prices; it is clear that computed prices are always very close relative to their level.

We also do the test for multiple equilibria from Corollary 4, and do not find evidence of multiple equilibria. So, although we do not know if the bargaining parameters in this model are sufficiently high for the results of Lemma 7 to apply, it appears they may be.

Table 4: Computation Results for Grennan’s (2013) Market for Stents

Panel A							
x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}
4.5407	4.5988	3.3802	3.5349	3.1868	3.4189	3.7342	3.2956
Panel B							
Reps.	Mean	St.D.	Min	25%	50%	75%	Max
100	0.446	12.31	0.144	0.144	0.144	0.144	0.8235

Notes: Panel A shows computed prices in thousands of dollars. The starting point is $(r_i + c_i)/2$ for each price. Each $X_{ij} \subset \mathbb{R}$ has 1,000 gridpoints, and the tolerance criterion for convergence is 10^{-5} . The first two stents are *DES*, the remaining are *BMS*. Panel B shows the results of restarting the computation at an additional 100, randomly selected initial price vectors. Each column shows a descriptive statistic of the distribution of L_∞ norms of the difference between the original and newly computed equilibria, in dollars. The largest difference is less than \$1, < 0.1% of the computed prices. Learning rate is $\alpha = 0.5$. The Jacobian determinant of $Q(\mathbf{x}; \mathbf{b})$ at the equilibrium is 0.9994. See Appendix D for details of the model.

9.3 Discontinuous Gains from Trade

We modify Horn and Wolinsky (1988)’s model so that retailers can buy from both suppliers. There are two suppliers and two retailers. Retailers compete in quantities, selling perfect substitutes with inverse demand $p = a - q_1 - q_2$. All firms have constant returns to scale. Suppliers’ marginal costs are normalized to zero and retailers’ marginal cost is the linear input price. Suppliers sell the same input and may sell to both retailers. U_i only sells to D_j if $x_{ij} \leq x_{-i,j}$. If $x_{ij} = x_{-i,j}$ we assume that each supplier satisfies half of D_j ’s demand.

Let x_j be the price at which D_j buys inputs. Under the linear inverse demand assumption, the equilibrium quantity supplied is $q_j = (a - 2x_j + x_{-j})/3$. Therefore, U_i ’s gains from trade are

$$\hat{\pi}_{U_i}(x_{ij}, \mathbf{x}_{-(ij)}) = \begin{cases} \sum_j \frac{x_{ij}(a-2x_{ij}+x_{-j})}{3} & x_{ij} < x_{-i,j} \forall j \\ \sum_j \frac{x_{ij}(a-2x_{ij}+x_{-j})}{6} & x_{ij} = x_{-i,j} \forall j \\ 0 & x_{ij} > x_{-i,j} \forall j \\ \vdots & \end{cases}$$

which is not continuous in $\mathbf{x}$. This model has a unique equilibrium with $x_{ij} = 0$ for all (ij) . This equilibrium has the undesirable property that oligopolistic suppliers have zero profits despite their bargaining power—the Bertrand paradox.

At first, the fact that eliminating the bilateral monopoly assumption results in the Bertrand paradox looks concerning. As Table 1 shows, however, empirical IO models that employ Nash-in-Nash use demand systems with differentiated products. In these models, inputs purchased from different suppliers

to produce differentiated goods are not perfect substitutes. In that case, adding bargaining pairs to the network will not create discontinuities in firms' gains from trade. Thus, in standard models the network of bargaining pairs is quite flexible.

10 Conclusion

We provide conditions for the existence and uniqueness of Nash-in-Nash equilibria, a bargaining equilibrium concept often used in applied IO. Existence and uniqueness matter for the estimation of structural models and comparative statics. To make our results useful for applied work, we prove that a wide class of models, which use discrete choice demand and constant returns to scale, satisfy our conditions. Our results apply to various types of contracts, do not assume symmetric firms, and do not depend on the number of firms, products, or the number of links in the bargaining network. We also provide an algorithm for the quick computation of Nash-in-Nash equilibria in models with multiple stages.

Our results can be extended to other settings, including longer supply chains, negotiations between unions and firms (González and Poblete, 2024), and tariff negotiations between countries (Bagwell et al., 2020, 2021). Alternatively, retailers in our model may be large final buyers.¹⁹

¹⁹Our conditions, however, are not substitutes for assumptions that are economically appropriate to the setting at hand. A specification could, for example, lead to the sum of a firm's gains from trade across all of its negotiations to add to more than its on-path profits. Such a property is undesirable and the model requires changes, but does not change the set of models to which our results apply.

Appendices

A Main Proofs

Proof of Lemma 1

Strategy of proof. We prove that under A2 the bargaining problem satisfies all the conditions of the maximum theorem. First we show that (1) N_{ij} is a continuous function of $(\mathbf{x}, b_{ij}) \in X \times (0, 1)$, (2) B_{ij} is a continuous correspondence of $(\mathbf{x}_{-(ij)}, b_{ij}) \in X_{-(ij)} \times (0, 1)$, and (3) B_{ij} is compact-valued and nonempty-valued. Then we show that the bargaining solution has all the properties stated in Lemma 1.

Notation. For readability, in the first part of the proof of Lemma 1 $x \in X_{-(ij)}$ stands for $\mathbf{x}_{-(ij)}$ and $y \in X_{ij}$ stands for $\mathbf{x}_{ij}$.

Proof of Lemma 1. *Step 1: Continuity of N_{ij} .* Continuous differentiability of gains from trade in $\mathbf{x} \in \mathcal{O} \subset \mathbb{R}^K$ (A2.1) implies their continuity in $\mathbf{x} \in X$. Continuity of the gains from trade in $\mathbf{x} \in X$ makes continuity in $(\mathbf{x}, b_{ij}) \in X \times (0, 1)$ trivial.

Step 2: Continuity of B_{ij} . *Claim A.1:* $PC_{ij}^{U_i}$ and $PC_{ij}^{D_j}$ are upper hemi-continuous (uhc) at every $\mathbf{x}_{-(ij)} \in X_{-(ij)}$. We prove uhc for $PC_{ij}^{U_i}$ only, since the proof for $PC_{ij}^{D_j}$ is identical.

By Theorem 3.4 of Stokey and Lucas Jr (1989), since $PC_{ij}^{U_i}$ is nonempty valued (A2.2: B_{ij} nonempty valued $\Rightarrow PC_{ij}^{U_i}$ nonempty valued) then, if the graph of $PC_{ij}^{U_i}$, $\text{graph} PC_{ij}^{U_i}$ is closed, $PC_{ij}^{U_i}$ is uhc. Note that

$$\text{graph} PC_{ij}^{U_i} = \{(y, x) : x \in X_{-(ij)} \text{ and } y \in PC_{ij}^{U_i}(x)\}. \quad (14)$$

Using (1) we can rewrite (14) as

$$\text{graph} PC_{ij}^{U_i} = \text{cl}_X(\{(y, x) \in X : \Delta\pi_{U_i}(y, x) > 0\}) \quad (15)$$

where cl_X is the closure of a set relative to X . The closure is closed, so the graph is closed. The same proof applies to $PC_{ij}^{D_j}$.

Claim A.2: B_{ij} is uhc. Again, all we need is that $\text{graph} B_{ij}$ be closed. The graph of the intersection of two nonempty-valued correspondences is the intersection of the graphs of those correspondences. That is

$$\text{graph} B_{ij} = \text{graph} PC_{ij}^{U_i} \cap \text{graph} PC_{ij}^{D_j}.$$

$\text{graph} PC_{ij}^{U_i}$ and $\text{graph} PC_{ij}^{D_j}$ are closed, so their intersection is closed.

Claim A.3. B_{ij} is an lower hemi-continuous (lhc) correspondence at every $\mathbf{x}_{-(ij)} \in X_{-(ij)}$.

By definition, B_{ij} is lhc at $x \in X_{-(ij)}$ if, for every open subset V of X_{ij} such that $V \cap B_{ij}(x) \neq \emptyset$, there exists $\delta > 0$ such that, for all x' in the open ball $\mathcal{N}_\delta(x) \subset X_{-(ij)}$, $B_{ij}(x') \cap V \neq \emptyset$.

We start by noting that, by A2.2, $B_{ij}(x)$ is a set with a non-empty interior. By A2.4, if y is in the boundary of X_{ij} , then $\Delta\pi_f(y, x) = 0$ for some $f \in \{U_i, D_j\}$. Therefore, for all $x \in X_{-(ij)}$, $\text{int}_{ij}(B_{ij}(x)) \subset \text{int}(X_{ij})$. (Below we prove that $\text{int}_{ij} = \text{int}$.)

Now, $y \in \text{int}_{ij}(B_{ij}(x))$ implies $y \in \text{int}(PC_{ij}^f(x))$ for all $f \in \{U_i, D_j\}$. Since V is open, there exists $y \in V \cap \text{int}_{ij}(B_{ij}(x))$. By continuity of each $\Delta\pi_f(y, x)$ in (y, x) , these functions are continuous in $x \in X_{-(ij)}$. So, for each f there exists $\delta_f > 0$ such that $d_{X_{-(ij)}}(x, x') < \delta_f$ implies $d_{\mathbb{R}}(\Delta\pi_f(y, x), \Delta\pi_f(y, x'))$ is so small that (1) $\text{int}(B(x')) \neq \emptyset$, and (2) $\Delta\pi_f(y, x') > 0$. But then $y \in PC_{ij}^f(x')$ for all $x' \in \mathcal{N}_{\delta_f}(x)$, so $PC_{ij}^f(x) \cap V \neq \emptyset$. Letting $\delta := \min\{\delta_{U_i}, \delta_{D_j}\}$ completes the proof of lhc at $\mathbf{x}_{-(ij)}$ for every $\mathbf{x}_{-(ij)} \in X_{-(ij)}$. Therefore, B_{ij} is lhc.

Since B_{ij} is uhc and lhc at all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, it is a continuous correspondence.

Step 3: B_{ij} is nonempty valued and compact valued. Nonempty-valued follows directly from A2.2. Compact-valued follows from the definitions (1) (closures are closed sets) and compactness of X in $\mathbb{R}^K$.

Step 4: Application of the maximum theorem. By the maximum theorem, for each $\mathbf{x}_{-(ij)} \in X_{-(ij)}$

$$\mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; b_{ij}) \in \arg \max_{\mathbf{x}_{ij}} \left\{ N_{ij}(\mathbf{x}_{ij}', \mathbf{x}_{-(ij)}; b_{ij}) : \mathbf{x}_{ij}' \in B_{ij}(\mathbf{x}_{-(ij)}) \right\} \quad (16)$$

is nonempty-valued, compact-valued, and uhc in $\mathbf{x}_{-(ij)}$. By A2.3 $\mathbf{x}_{ij}^*$ is singleton-valued, so it is a continuous function.

Step 5: Additional properties. Invariance to positive affine transformations of profits is trivial, as they induce monotone, continuous transformations of the objective function. For Pareto efficiency, if there is an $\mathbf{x}_{ij}'$ that is a Pareto improvement on $\mathbf{x}_{ij}^*$, then $\mathbf{x}_{ij}^*$ is not a global maximizer—a contradiction. For independence of irrelevant alternatives: since $\mathbf{x}_{-(ij)}$ is the same and $\emptyset \neq B' \subsetneq B_{ij}(\mathbf{x}_{-(ij)})$, the set of global maximizers of $N_{ij}(\cdot, \mathbf{x}_{-(ij)}; b_{ij})$ in B' is the same as in $B_{ij}(\mathbf{x}_{-(ij)})$.

If $b_{ij} \in (0, 1)$ the bargaining solution must be jointly profitable. To see this, assume it is not jointly profitable. Then, $N_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; b_{ij}) = 0$. By A2.3 there exists $\mathbf{x}_{ij}' \in B_{ij}(\mathbf{x}_{-(ij)})$ such that $N_{ij}(\mathbf{x}_{ij}', \mathbf{x}_{-(ij)}; b_{ij}) > 0$, so $\mathbf{x}_{ij}^*$ is not the bargaining solution. By A2.4 any $\mathbf{x}_{ij}$ in the boundary of X_{ij} is not jointly profitable, so the interior of $B_{ij}(\mathbf{x}_{-(ij)})$ relative to X_{ij} is always in the interior of X_{ij} relative to $\mathbb{R}^{K_{ij}}$. Thus, the interior of $B_{ij}(\mathbf{x}_{-(ij)})$ relative to X_{ij} coincides with the interior of $B_{ij}(\mathbf{x}_{-(ij)})$ relative to $\mathbb{R}^{K_{ij}}$ for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$. Therefore, if $b_{ij} \in (0, 1)$ the bargaining solution function maps into the interior of X_{ij} (relative to $\mathbb{R}^K$). (This proves the claim above and in footnote 5 that $\text{int}_{ij} = \text{int}$.) ■

A.1 Proof of Corollary 1

The proof of Lemma 1 only uses continuity of the gains from trade in $\mathbf{x}_{ij}$, which is implied by continuous differentiability. ■

Proof of Proposition 1

Given $\mathbf{b}$, by Lemma 1 F is a continuous function from X into its interior. X is compact and convex, so by the Brouwer fixed-point theorem it has a fixed point. ■

A.2 Proof of Corollary 2

Direct from Corollary 1 and the proof of Proposition 1.

Proof of Proposition 2

Strategy of proof. We use the Poincaré-Hopf and transversality theorems to show the number of fixed points is finite for Lebesgue-almost every $\mathbf{b} \in (0, 1)^{|G|}$. The Poincaré-Hopf theorem (Milnor, 1965; Simsek et al., 2007; Morrow and Skerlos, 2010b) gives the conditions for finiteness. We use the transversality theorem (Mas-Colell et al., 1995) to show these conditions hold for Lebesgue-almost every $\mathbf{b} \in (0, 1)^{|G|}$.

Preliminaries. Consider function $Q : X \times (0, 1)^{|G|} \rightarrow \mathbb{R}^K$. Q is a function of $\mathbf{x}$ parameterized by $\mathbf{b}$ such that

$$Q(\mathbf{x}; \mathbf{b}) = \mathbf{x} - F(\mathbf{x}, \mathbf{b}). \quad (17)$$

A Nash-in-Nash equilibrium is a solution of $Q(\mathbf{x}, \mathbf{b})=0$. Let E be the set of solutions. To use the Poincaré-Hopf theorem the following must hold:

1. Q is a continuously differentiable (i.e. C^1) vector field. This holds by A3 and the definition of Q .
2. The Euler characteristic of X must be 1: $\chi(X) = 1$. This holds because X is convex.
3. Q must **point outwards at the boundary** of X . That is, if $\mathbf{x} \in \partial X$, then there exists a sequence $\varepsilon_n \downarrow 0$ such that

$$\mathbf{x} + \varepsilon_n Q(\mathbf{x}; \mathbf{b}) \notin X \forall n \in \mathbb{N}. \quad (18)$$

4. The Jacobian of Q at V is non-singular. That is, for all $\mathbf{x} \in V$

$$\det(I - \nabla_{\mathbf{x}} F(\mathbf{x}; \mathbf{b})) \neq 0 \quad (19)$$

The Poincaré-Hopf theorem states that, if the conditions above hold, then

$$\sum_{\mathbf{x} \in E} \text{sign}(\det(\nabla_{\mathbf{x}} Q(\mathbf{x}; \mathbf{b}))) = \chi(X) \quad (20)$$

(see Simsek et al. (2007)). Since $\chi(X) = 1$ in our setting, this means that, if all the conditions above hold, there is a finite, odd number of isolated (and so, locally unique) equilibria. Thus, it remains to show that (18) and (19) hold for almost every $\mathbf{b} \in (0, 1)^{|G|}$.

Step 1. Q points outwards at the boundary. Fix any $\mathbf{b} \in (0, 1)^{|G|}$. Let $c \in \{1, \dots, K\}$ index the dimensions of X . if $\mathbf{x} \in \partial X$, then there exists c such that $x_c \in \mathbb{R}$ (the element of $\mathbf{x}$ in the c -th dimension) is in ∂X_c . Because X is convex, one of the following cases must hold.

Case 1: For all $z > x_c$, $z \notin X_c$. Because F maps into $\text{int}(X)$, $F_c(\mathbf{x}; \mathbf{b}) < x_c$. Therefore

$$\begin{aligned} x_c + \varepsilon Q_c(\mathbf{x}; \mathbf{b}) &> x_c \quad \forall \varepsilon > 0 \\ \iff \varepsilon(x_c - F_c(\mathbf{x}; \mathbf{b})) &> 0 \quad \forall \varepsilon > 0. \end{aligned}$$

But the second line holds because $\varepsilon > 0$ and $x_c > F_c(\mathbf{x}; \mathbf{b})$ by assumption.

Case 2: For all $z < x_c$, $z \notin X_c$. Because F maps into $\text{int}(X)$, $F_c(\mathbf{x}; \mathbf{b}) > x_c$. Therefore

$$\begin{aligned} x_c + \varepsilon Q_c(\mathbf{x}; \mathbf{b}) &< x_c \quad \forall \varepsilon > 0 \\ \iff \varepsilon(x_c - F_c(\mathbf{x}; \mathbf{b})) &< 0. \end{aligned}$$

But, again, the second line holds because $\varepsilon > 0$ and $x_c < F_c(\mathbf{x}; \mathbf{b})$ by assumption. Therefore $Q(\cdot; \mathbf{b})$ points outwards at the boundary for every $\mathbf{b} \in (0, 1)^{|G|}$.

Step 2. Jacobian is nonsingular at $\mathbf{x} \in E$. We prove this holds for Lebesgue-almost every $\mathbf{b} \in (0, 1)^{|G|}$ using the transversality theorem. For clarity, assign to each bargaining pair (ij) a number $g \in \{1, \dots, |G|\}$. The vector of partial derivatives of the bargaining solution of pair g , $\mathbf{x}_g^*$ with respect to $b_{g'}$ is $\partial \mathbf{x}_g^* / \partial b_{g'}$.

Construct the $K \times |G|$ matrix $D_{\mathbf{b}}F(\mathbf{x}; \mathbf{b})$ as follows: in the first column, the first K_1 rows correspond to the vector $\partial \mathbf{x}_{g=1}^* / \partial b_{g=1}$, for (ij) such that $g = 1$. In the first $K_{ij} = K_1$ rows, the remaining $|G| - 1$ columns correspond to $\partial \mathbf{x}_g^* / \partial b_{g'}$, $g' > 1$, ordered by $g' \in \{2, \dots, |G|\}$. In the first column, rows $K_1 + 1$ to K_2 correspond to $\partial \mathbf{x}_{g=2}^* / \partial b_{g=1}$, and so on.

By A3.2, matrix $D_{\mathbf{b}}F(\mathbf{x}; \mathbf{b})$ is block diagonal with nonzero entries in the diagonal blocks: the blocks in the diagonal are the vectors $\partial \mathbf{x}_{g=1}^* / \partial b_{g=1}$, and all other entries are zero.

To finish this step we use the concept of a regular system (Mas-Colell et al., 1995) and the transversality theorem, as stated in Mas-Colell et al. (1995) (theorem M.E.2). The definition of regularity and the theorem are restated here. For our purposes, let $A = \text{int}(X)$ and $B = (0, 1)^{|G|}$.

Definition A.1: Regularity. Let $A \subset \mathbb{R}^K$ and $B \subset \mathbb{R}^{|G|}$ be open sets, and let the system of equations $Q(\cdot; \mathbf{b})$ be continuously differentiable. We say the system is regular at $\mathbf{b} \in B$ if, at any solution $\mathbf{x}$, the matrix of derivatives with respect to $\mathbf{x}$, $D_{\mathbf{x}}Q(\mathbf{x}; \mathbf{b})$, is invertible.

Transversality Theorem. Suppose we are given open sets $A \subset \mathbb{R}^K$ and $B \subset \mathbb{R}^{|G|}$, and a continuously differentiable function $Q : A \times B \rightarrow \mathbb{R}^K$. If $Q(\cdot; \cdot)$ is such that the $K \times (K + |G|)$ matrix

$$\begin{bmatrix} \underbrace{D_{\mathbf{x}}Q(\mathbf{x}, \mathbf{b})}_{K \times K} & \underbrace{D_{\mathbf{b}}Q(\mathbf{x}, \mathbf{b})}_{K \times |G|} \end{bmatrix} \quad (21)$$

has row rank K whenever $Q(\mathbf{x}, \mathbf{b}) = 0$, then the system with K equations and K unknowns is regular for almost every $\mathbf{b}$ in B .

It is clear that we want to prove regularity of $Q(\cdot; \mathbf{b})$. The transversality theorem lets us do that for Lebesgue-almost every $\mathbf{b} \in (0, 1)^{|G|}$. By A3.1, the continuous differentiability condition of the transversality theorem holds. Thus, it remains to prove the invertibility of (21). Because $Q(\mathbf{x}; \mathbf{b}) = \mathbf{x} - F(\mathbf{x}; \mathbf{b})$, we can rewrite (21) as

$$C = [I_K - D_{\mathbf{x}}F(\mathbf{x}, \mathbf{b}) \mid -D_{\mathbf{b}}F(\mathbf{x}, \mathbf{b})]$$

where I_K is the $K \times K$ identity matrix.

Case 1: $K = |G|$. Then $K_{ij} = 1$ for all (ij) , so by A3.2 $D_{\mathbf{b}}F(\mathbf{x}, \mathbf{b})$ is a diagonal matrix with row rank K . Then C has row rank K and the system is regular.

Case 2: $K > |G|$. Then, there exists at least one (ij) such that $K_{ij} > 1$. Assume, by contradiction, that (21) has row rank smaller than K . Let $r \in \{1, \dots, K\}$ index a linearly dependent row, and let $\bar{g}$ be the pair to which the contract term whose derivatives are linearly dependent belongs. By the description of $-D_{\mathbf{b}}F(\mathbf{x}, \mathbf{b})$ above, this means that the rows whose linear combination replicate row r are rows of derivatives of the other contract terms of pair $\bar{g}$. Without loss of generality, let $\bar{g} = 1$. Then, the first K_{ij} columns of those K_{ij} rows form a $K_{ij} \times K_{ij}$ identity matrix, which has full row rank. This contradicts r being linearly dependent. Therefore, (21) has full row rank.

Since (21) has full row rank, by the transversality theorem the system is regular for almost every $\mathbf{b} \in (0, 1)^{|G|}$, meaning (19) holds for Lebesgue-almost every $\mathbf{b} \in (0, 1)^{|G|}$. Then, by the Poincaré-Hopf theorem, for Lebesgue-almost every $\mathbf{b} \in (0, 1)^{|G|}$ there is a finite number of (isolated) solutions for $Q(\mathbf{x}, \mathbf{b}) = 0$. $\mathbf{x}$ is a Nash-in-Nash equilibrium if, and only if, it is a solution of $Q(\mathbf{x}, \mathbf{b}) = 0$, completing the proof. ■

Proof of Proposition 3

Again we use the Poincaré-Hopf theorem. Continuous differentiability, $\chi(X) = 1$, and (18) hold by the same arguments as the proof of Proposition 2. (19) holds by A4 (positive Jacobian determinant at every Nash-in-Nash equilibrium), so we do not need A3.2 for this proof. Since the Jacobian determinant is always positive, the index is always 1. So, for (20) to hold, E must be a singleton. Since all the conditions of the Poincaré-Hopf theorem hold, (20) holds, completing the proof. ■

Proof of Lemma 2

The proof has two parts. First, we prove a simple relation between the eigenvalues of $D_{\mathbf{x}}F$ and $D_{\mathbf{x}}Q$. Then we use Gershgorin's Circle Theorem to prove that the hypothesis of the lemma implies A4.

Step 1: Relation between the eigenvalues of $D_{\mathbf{x}}F$ and $D_{\mathbf{x}}Q$. Let λ_c and μ_c be the eigenvalues of $D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b})$ and $D_{\mathbf{x}}Q(\mathbf{x}; \mathbf{b})$, respectively, with $c \in \{1, \dots, K\}$ (so these eigenvalues include multiplicity). I and $D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b})$ commute (i.e. $ID_{\mathbf{x}}F(\mathbf{x}; \mathbf{b}) = D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b})I$) so, by Schur's triangularization theorem, there exists a permutation of eigenvalues such that

$$\mu_c = 1 - \lambda_c \quad \forall c \quad (22)$$

(Horn and Johnson, 2013, Theorem 2.4.8.1, p. 117).

Step 2. Let ρ denote the spectral radius of a matrix. Since complex eigenvalues come in conjugate pairs, a sufficient condition for the determinant of $D_{\mathbf{x}}Q(\mathbf{x}; \mathbf{b})$ to be positive is that no real eigenvalues are larger than or equal to one. Using (3) and (4), and the upper bound on the spectrum given by

Gershgorin's Circle Theorem, we have

$$\rho(D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b})) \leq \min \left\{ \max_{k \in \mathcal{K}_{ij} : (ij) \in G} \left\{ \sum_{(i'j') \neq (ij)} \sum_{k' \in \mathcal{K}_{(i'j')}} \left| \frac{\partial x_{i'j'k'}}{\partial x_{ijk}} \right| \right\}, \max_{k \in \mathcal{K}_{ij} : (ij) \in G} \left\{ \sum_{(i'j') \neq (ij)} \sum_{k' \in \mathcal{K}_{(i'j')}} \left| \frac{\partial x_{ijk}}{\partial x_{i'j'k'}} \right| \right\} \right\} < 1,$$

(Horn and Johnson, 2013, Corollary 6.1.5, p. 390). So, all the real eigenvalues of $D_{\mathbf{x}}Q(\mathbf{x}; \mathbf{b})$ are positive, and its determinant is too. ■

Lemma A.1

Lemma A.1 shows that stability (the conditions of Lemma 2) is not necessary for global uniqueness. It does this by showing that global uniqueness can obtain even if the spectral radius of F is greater than 1, which would violate stability.

Lemma A.1. *If any of the following hold, Assumption 4 holds: For every fixed point $\mathbf{x}$*

1. *The spectral radius of $D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b})$ is smaller than 1.*
2. *The absolute value of spectral abscissa of $D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b})$ is smaller than 1.²⁰*
3. *The number of real eigenvalues of $D_{\mathbf{x}}F(\mathbf{x}; \mathbf{b})$ that are greater or equal than 1 is even.*

Proof of Lemma A.1. *Part 1.* Follows directly from the proof of Lemma 2.

Part 2. Since complex eigenvalues come in conjugate pairs, a sufficient condition for the determinant of $D_{\mathbf{x}}Q(\mathbf{x}; \mathbf{b})$ to be positive is that no real eigenvalues are larger than or equal to one. This is true if the absolute value of the spectral abscissa is smaller than one.

Part 3. The determinant of a matrix is the product of its eigenvalues. By (22), the determinant of $D_{\mathbf{x}}Q(\mathbf{x}, \mathbf{b})$ is positive if an even number (including zero) of eigenvalues is negative. That is true if an even number of eigenvalues of $D_{\mathbf{x}}F(\mathbf{x}, \mathbf{b})$ is greater than or equal to one. In practice, the proof of Proposition 2 shows we can ignore the equality to one. ■

Comment. Part 3 of Lemma A.1 is a necessary and sufficient condition for Assumption 4, but is generally hard to check without knowledge of F . The conditions in part 1 of Lemma A.1 are the least general - they imply that parts 2 and 3 hold -, but it is the easiest to check using spectral analysis. The conditions in part 2 imply that of part 3.

The conditions in Lemma A.1 mean that stability is not necessary because stability, as understood here (Moulin, 1984), implies all real eigenvalues of the Jacobian are smaller than 1 in absolute value. As Lemma A.1 shows, this condition is not necessary in the class of models studied in this paper.

²⁰The spectral abscissa of a matrix is the greatest of the real parts of its eigenvalues.

Proof of Corollary 3

If contracts are efficient then, by the definition given in the introduction, given $\mathbf{b} \in (0, 1)^{|G|}$ and for all $(ij) \in G$, the bargaining solution $\mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}, b_{ij})$ is constant with respect to $\mathbf{x}_{-(ij)}$. Thus there is a globally unique equilibrium. ■

Proof of Corollary 4

A2 and A3.1 hold, so by the proof of Proposition 2 expression (20) holds. However, the nonpositive determinant condition of Corollary 4 means that (20) can only hold if E is not a singleton. ■

Proof of Proposition 4

Strategy of proof. First we prove that retailers' equilibrium strategies in the second stage are three times continuously differentiable in contracts. The chain rule finishes the proof. The proof that on-path strategies are three times continuously differentiable in $\mathbf{x}$ is the same as the proof that off-path strategies are three times continuously differentiable in $\mathbf{x}_{-(ij)}$, so we only prove the result on-path.

Step 1: Equilibrium strategies are three times continuously differentiable in contract terms. By A5, each π_{D_j} is at least four times continuously differentiable in $\mathbf{p}_j$. By A1 and A5.2, each D_j 's FOC with respect to $\mathbf{p}_j$ holds in the second stage equilibrium $\mathbf{p}^* \equiv (\mathbf{p}_j^*, \mathbf{p}_{-j}^*)$:

$$D_{\mathbf{p}_j} \pi_{D_j}(\mathbf{p}_j^*, \mathbf{p}_{-j}^*, \mathbf{x}) = 0 \quad \forall j \in \mathcal{J}$$

We collect all these FOCs in the second stage equilibrium system, described in the main body of the paper. By A5.1, this system is three continuously differentiable in $\mathbf{p}$, and its Jacobian with respect to $\mathbf{p}$ is invertible (A5.2). Therefore, by the implicit function theorem we obtain the matrix of first partial derivatives of $\mathbf{p}^*(\mathbf{x})$ with respect to $\mathbf{x}$:

$$D_{\mathbf{x}} \mathbf{p}^*(\mathbf{x}) = - \left[D_{\mathbf{p}} [D_{\mathbf{p}_j} \pi_{D_j}(\mathbf{p}_j^*, \mathbf{p}_{-j}^*)]_j \right]^{-1} \left[D_{\mathbf{x}} [D_{\mathbf{p}_j} \pi_{D_j}(\mathbf{p}_j^*, \mathbf{p}_{-j}^*)]_j \right]$$

The existence of the arrays of (continuous) second and third partial derivatives of $\mathbf{p}^*(\mathbf{x})$ with respect to $\mathbf{x}$ follows from A5.1 and the rules of matrix differentiation applied to the inverse of a matrix. In particular, if $A(\mathbf{x})$ is a square matrix whose entries are smooth functions of vector $\mathbf{x}$, then

$$\frac{\partial A^{-1}}{\partial \mathbf{x}} = A^{-1} \frac{\partial A}{\partial \mathbf{x}} A^{-1}$$

Step 2: Application of the chain rule. Consider pair (ij) and firm $f \in \{U_i, D_j\}$. We can write the on-path profits function as $\hat{\pi}_f(\mathbf{p}^*(\mathbf{x}), \mathbf{x})$. A5.1 gives three times continuous differentiability with respect to $\mathbf{x}$. Repeated application of the chain rule, and existence of the arrays of (continuous) second and third partial derivatives of $\mathbf{p}^*(\mathbf{x})$ with respect to $\mathbf{x}$ complete the proof. ■

Proof of Corollary 5

Relaxing A5.1 to three times continuous differentiability, and following the exact same steps as the proof of Proposition 4 (save for getting the array of third partial derivatives) delivers the proof for twice continuous differentiability. ■

Proof of Lemma 4

We must consider three cases. Case 1 is when a breakdown of negotiations between U_i and D_j means D_j can no longer sell one or more products in the retail market. Case 2 is when breakdown of negotiations forces D_j to sell at higher costs (e.g., because it must now buy inputs from an inefficient outside option, or produce in-house). Case 3 is a combination of both.

Let $0 < y < \infty$ be the income (or wealth) of the consumer with the highest income. In models with mixed logit demand income can vary by consumers; in other models it cannot. The sum of all consumers' income/wealth is $\bar{y}$. $\bar{y}$ is an upper bound on total industry revenue.

The set of products retailer D_j sells is L_j . The set of all products in the retail market is L . The collection of sets L_j forms a partition of L . The market share of product l is $s_l(\mathbf{p})$, which we also write as $s_l(\mathbf{x})$ to save notation. Without loss of generality we normalize the size of the market to one, and assume that a consumer indifferent between a product in L or the outside good chooses the outside good. A retailer's marginal cost for l is mc_l .

1. *Linear input prices.* For each $(ij) \in G$, let $X_{ij} = [-\bar{y}, \bar{y}]^{K_{ij}}$. It follows from Lemma 3 that A2.4 is satisfied. Now we consider the three cases outlined above,

Case 1: Buying from U_i lets D_j sell more products. Let $L_{ij} \subset L_j$ be the set of products for whose production D_j uses inputs purchased from U_i . D_j 's gains from trade are

$$\begin{aligned} \Delta\pi_{D_j}(\mathbf{x}) = & \underbrace{\sum_{l \in L_{ij}} (p_l(\mathbf{x}) - mc_l(\mathbf{x}_{ij}))s_l(\mathbf{x}) + \sum_{l \in L_j \setminus L_{ij}} (p_l(\mathbf{x}) - mc_l(\mathbf{x}_{-(ij)}))s_l(\mathbf{x})}_{\hat{\pi}_{D_j}(\mathbf{x})} \\ & - \underbrace{\sum_{l \in L_j \setminus L_{ij}} (p_l(\mathbf{x}_{-(ij)}) - mc_l(\mathbf{x}_{-(ij)}))s_l(\mathbf{x}_{-(ij)})}_{\bar{\pi}_{D_j}(\mathbf{x}_{-(ij)})} \end{aligned} \quad (23)$$

With linear input prices, c_l , $l \in L_{ij}$ is strictly increasing in each x_{ijk} , $k = 1, \dots, K_{ij}$. It follows that for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, there exists at least one $\mathbf{x}$ such that $c_l = y$ for all $l \in L_{ij}$. Non-negativity of markups means the first summation in (23) becomes zero. Since all products in L_{ij} are priced out of the market, the second and third summations are equal, and D_j 's gains from trade are zero. This means that U_i sells nothing to D_j , while all its other sales to other retailers (and the unit prices) are the same on and off-path. Therefore, U_i also has zero gains from trade.

Case 2. Buying from U_i lets D_j sell the same products at lower cost. Let $\bar{c}_{ijk}$ be the price at which D_j can buy input $k = 1, \dots, K_{ij}$ if negotiations break down. U_i 's marginal cost for that product is $c_{ijk} < \bar{c}_{ijk}$. D_j 's gains from trade are

$$\begin{aligned} \Delta\pi_{D_j}(\mathbf{x}) = & \underbrace{\sum_{l \in L_{ij}} (p_l(\mathbf{x}) - mc_l(\mathbf{x}_{ij}))s_l(\mathbf{x}) + \sum_{l \in L_j \setminus L_{ij}} (p_l(\mathbf{x}) - mc_l(\mathbf{x}_{-(ij)}))s_l(\mathbf{x}) -}_{\hat{\pi}_{D_j}(\mathbf{x})} \\ & \underbrace{\sum_{l \in L_{ij}} (p_l(\bar{\mathbf{c}}_{ij}, \mathbf{x}_{-(ij)}) - mc_l(\bar{\mathbf{c}}_{ij}))s_l(\bar{\mathbf{c}}_{ij}, \mathbf{x}_{-(ij)}) + \sum_{l \in L_j \setminus L_{ij}} (p_l(\bar{\mathbf{c}}_{ij}, \mathbf{x}_{-(ij)}) - mc_l(\mathbf{x}_{-(ij)}))s_l(\mathbf{x}_{-(ij)})}_{\bar{\pi}_{D_j}(\bar{\mathbf{c}}_{ij}, \mathbf{x}_{-(ij)})} \end{aligned}$$

Therefore, setting $\mathbf{x}_{ij} = \bar{\mathbf{c}}_{ij}$ gives D_j zero gains from trade. Since $c_{ijk} < \bar{c}_{ijk}$ for all $k = 1, \dots, K_{ij}$, $\mathbf{x}_{ij} = \bar{\mathbf{c}}_{ij}$ gives U_i nonnegative markups in each input it sells, and non-negative gains from trade.

Case 3: Combination of cases 1 and 2. Combine the proofs above.

2. Two-part tariffs. Set the lump-sum transfer to zero. Then, for the unit prices, use the same proof as for linear input prices.

3. Revenue sharing. The proof follows the three cases used in linear input prices.

Case 1: Buying from U_i lets D_j sell more products. An increase in revenue share x_{ijk} increases the marginal cost of every product input k is used to produce. Therefore, there exists a contract $\mathbf{x}_{ij}$ such that $mc_l = y$ for all $l \in L_{ij}$. Just as for linear input prices, non-negativity of markups completes the proof of case 1.

Case 2. Buying from U_i lets D_j sell the same products at lower cost. As in the proof of Lemma 3, for each k , set x_{ijk} such that the revenue transferred to U_i is $\bar{c}_{ijk}$ per unit of input. This makes the contract equivalent to the outside option for D_j , and gives U_i positive gains from trade.

Case 3: Combination of cases 1 and 2. Again, combine the proofs above. ■

Proof of Proposition 5

By twice continuous differentiability of the gains from trade (A2.1), N_{ij} is twice continuously differentiable in $(\mathbf{x}, b_{ij})$ if $\mathbf{x}_{ij} \in \text{int}(B_{ij}(\mathbf{x}_{-(ij)}))$. A2 holds, so by Lemma 1 the bargaining solutions are always in the interior of X . By A2.2, the bargaining solution is always in the interior of the bargaining set (see the proof of Lemma 1). Therefore, for all $(ij) \in G$, for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, the following FOC holds at the bargaining solution:

$$D_{\mathbf{x}_{ij}} N_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; b_{ij}) = 0 \quad (24)$$

By the implicit function theorem, we can obtain the (continuous) partial derivatives with respect to $(\mathbf{x}_{-(ij)}, b_{ij})$ with

$$D_{(\mathbf{x}_{-(ij)}, b_{ij})} \mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; b_{ij}) = - \left[D_{\mathbf{x}_{ij} \mathbf{x}_{ij}}^2 N_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; b_{ij}) \right]^{-1} D_{\mathbf{x}_{ij}(\mathbf{x}_{-(ij)}, b_{ij})}^2 N_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; b_{ij}). \quad (25)$$

By A7 the Hessian of $N_{ij}(\cdot, \mathbf{x}_{-(ij)}; b_{ij})$ is always invertible at the bargaining solution, completing the proof. For the second partial derivatives, we use the chain rule and the rule of matrix differentiation for inverted matrices. ■

Proof of Proposition 6

Notation. $\bar{c} \in \mathbb{R}_{++}$ is the upper boundary on marginal costs of each traded input. The upper bound is the same for each input. If a retailer sells the same input to more than one supplier, we count the input once for each supplier. Vector $\mathbf{c}_{ij} \in [0, \bar{c}]^{K_{ij}} \subset \mathbb{R}^{K_{ij}}$ is the vector of marginal costs of production of the inputs that U_i sells to D_j . Vector $\mathbf{c} \in [0, \bar{c}]^K \subset \mathbb{R}^K$ is the vector of supplier marginal costs for all the inputs traded by all bargaining sets.

Because $[0, \bar{c}] \setminus (0, \bar{c})$ has Lebesgue-measure zero in $\mathbb{R}^K$, if a subset of $(0, \bar{c})$ has Lebesgue-measure zero in $\mathbb{R}^K$ then it is also a Lebesgue-measure zero subset of $[0, \bar{c}]$. If a set has Lebesgue measure zero in $\mathbb{R}^{K_{ij}}$, then it has Lebesgue-measure zero in $\mathbb{R}^K$.

For the functional form of firms' on and off-path profits we use the notation of Section 9.1, because it generalizes to all models with discrete choice demand and constant returns to scale.

Pick any $(ij) \in G$ and any $\mathbf{x}_{-(ij)} \in X_{-(ij)}$. We begin by using the FOC of the bargaining problem. From the proof of Proposition 5 we know that (24) holds at the bargaining solution. We use the full expression for $\Delta\pi_{U_i}(\mathbf{x}_{-(ij)})$ in expression (12), since it generalizes to all models with discrete choice demand and constant returns to scale. For each input $h \in H_i$ that U_i sells to D_j , define

$$q_{jh}(\mathbf{x}) \equiv \sum_{l \in L_j} a_{jhl} s_l(\mathbf{x}). \quad (26)$$

Then, letting $h \in \mathcal{K}_{ij}$ we can write U_i 's gains from trade as

$$\Delta\pi_{U_i}(\mathbf{x}) = \sum_{h \in H_i} q_{jh}(\mathbf{x})(x_h - c_h) + \underbrace{\sum_{j' \neq j: (ij') \in G} \sum_{h \in H_i} q_{j'h}(\mathbf{x})(x_h - c_h)}_{\equiv \tilde{\pi}_{U_i}(\mathbf{x})} - \underbrace{\sum_{j' \neq j: (ij') \in G} \sum_{h \in H_i} q_{j'h}(\mathbf{x}_{-(ij)})(x_h - c_h)}_{\equiv \bar{\pi}_{U_i}(\mathbf{x})} \quad (27)$$

We can write FOC of the bargaining problem w.r.t. $x_h, h \in \mathcal{K}_{ij}$ as

$$\underbrace{b_{ij} \frac{\frac{\partial \Delta\pi_{D_j}}{\partial x_h}}{\Delta\pi_{D_j}(\mathbf{x})}}_{<0} + (1 - b_{ij}) \frac{q_{jh} + \sum_{k \in H_i} \frac{\partial q_{jk}}{\partial x_h}(\mathbf{x})(x_k - c_k) + \tilde{\pi}'_{U_i}(\mathbf{x})}{\sum_{h \in H_i} q_{jh}(\mathbf{x})(x_h - c_h) + \tilde{\pi}_{U_i}(\mathbf{x}) - \bar{\pi}_{U_i}(\mathbf{x})} = 0 \quad (28)$$

where $\tilde{\pi}'_{U_i, h}(\mathbf{x})$ is shorthand for the partial derivative of $\tilde{\pi}_{U_i}(\mathbf{x})$ w.r.t. x_h . Collecting all $q_{jk}(\mathbf{x})$ in K_{ij} -dimensional vector $\mathbf{q}$, all $\partial q_{jk}(\mathbf{x})/\partial x_h$ in K_{ij} -dimensional vector $\partial_h \mathbf{q}$, and all $(x_h - c_h), h \in \mathcal{K}_{ij}$ in K_{ij} -dimensional vector $(\mathbf{x}_{ij} - \mathbf{c}_{ij})$ we can rewrite (28) as

$$\frac{b_{ij}}{1 - b_{ij}} \frac{A}{B} [\mathbf{q}^T (\mathbf{x} - \mathbf{c}_{ij}) + \tilde{\pi}_{U_i}(\mathbf{x}) - \bar{\pi}_{U_i}(\mathbf{x})] + \mathbf{q} + \partial_h \mathbf{q}^T (\mathbf{x} - \mathbf{c}_{ij}) + \tilde{\pi}'_{U_i, h}(\mathbf{x}) = 0 \quad (29)$$

with $A = \partial \Delta\pi_{D_j} / \partial x_h < 0$ and $B = \Delta\pi_{D_j}(\mathbf{x}) > 0$ (A2.2). This follows the notation of Gowrisankaran et al. (2015). Separating all the terms multiplied by $(\mathbf{x}_{ij} - \mathbf{c}_{ij})$ and stacking all K_{ij} FOCs we can rewrite the Jacobian at the bargaining solution as

$$(\Omega + \Lambda)(\mathbf{x}_{ij} - \mathbf{c}_{ij}) + \mathbf{q} + \Delta = 0$$

where $K_{ij} \times K_{ij}$ matrices Ω and Λ come from stacking the vectors in 29 and Δ is a $K_{ij} \times 1$ column vector that collects all other terms. From Gowrisankaran et al. (2015) we know that $(\Omega + \Lambda)$ is invertible, so we get that the bargaining solution is one of the solutions to the following system of nonlinear equations

$$V_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) = \mathbf{x}_{ij} - \mathbf{c}_{ij} + (\Omega + \Lambda)^{-1}(\mathbf{q} + \Delta) = 0 \quad (30)$$

Note that, given $(\mathbf{x}_{-(ij)}; b_{ij}, \mathbf{c}_{ij})$, we obtain $V_{ij}(\cdot, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})$ by multiplying each element of the FOC of the bargaining problem, (28), by a non-zero real valued function $g_{ij}(\cdot, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})$. (Here we make the dependence of functions on $\mathbf{c}_{ij}$ explicit, leaving dependence on the remaining elements of vector $\mathbf{c}$ implicit.) The diagonal matrix that collects these functions in the diagonal is $G_{ij}(\cdot, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})$. By Assumption 2.2, the interior bargaining set $B_{ij}(\mathbf{x}_{-(ij)})$ is nonempty, so this transformation always exists (at least locally around $\mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})$). That is

$$V_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) = G_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) D_{\mathbf{x}_{ij}} N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}). \quad (31)$$

We want the bargaining solution (one of the solutions of (30)) to be continuously differentiable. As in the proof of Proposition 4, we can use the implicit function theorem *if* the system is regular at the bargaining solution, so proving regularity completes that part of the proof. However, Ω , Λ , $\mathbf{q}$, and Δ are nonlinear functions of $\mathbf{x}_{ij}$ (given $\mathbf{x}_{-(ij)}$, $\mathbf{c}$, and b_{ij}), so proving that the Jacobian determinant (i.e., determinant of the Jacobian) of (30) is nonzero is in general not feasible. Therefore, we use the transversality theorem again. We first want to prove that

$$[D_{\mathbf{x}_{ij}} V_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) | D_{\mathbf{c}_{ij}} V_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})] \quad (32)$$

has full row rank. But it is trivial from (30) that $D_{\mathbf{c}_{ij}} V_{ij}(\mathbf{x}_{ij}^*; \mathbf{c}_{ij}, \mathbf{x}_{-(ij)}) = -I_{K_{ij}}$ for all $(\mathbf{c}, \mathbf{x}_{-(ij)}, \mathbf{b}) \in (0, \bar{c}) \times X_{-(ij)} \times (0, 1)^{|G|}$, so we have full row rank. Therefore, by the transversality theorem, for Lebesgue-almost every $\mathbf{c}_{ij} \in (0, \bar{c})^{K_{ij}}$ (and so, for Lebesgue-almost every $\mathbf{c} \in (0, \bar{c})^K$), $V_{ij}(\cdot, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})$ is a regular system in $B_{ij}(\mathbf{x}_{-(ij)}) \times (0, \bar{c})^{K_{ij}}$.

To map regularity of $V_{ij}(\cdot, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})$ into regularity of $D_{\mathbf{x}_{ij}} N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})$, note that by the product rule we have

$$\begin{aligned} D_{\mathbf{x}_{ij}} V_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) &= D_{\mathbf{x}_{ij}} G_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) \underbrace{D_{\mathbf{x}_{ij}} N_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})}_{=0} \\ &\quad + G_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) D_{\mathbf{x}_{ij} \mathbf{x}_{ij}}^2 N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) \end{aligned}$$

Therefore

$$\begin{aligned} \det \left(D_{\mathbf{x}_{ij}} V_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) \right) &= \det \left(G_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) D_{\mathbf{x}_{ij} \mathbf{x}_{ij}}^2 N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) \right) \\ &= \det \left(G_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) \right) \\ &\quad \cdot \det \left(D_{\mathbf{x}_{ij} \mathbf{x}_{ij}}^2 N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) \right). \end{aligned}$$

Since $G_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}, b_{ij})$ is diagonal with non-zero elements in its diagonal, it has non-zero determi-

nant. So, given $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, for Lebesgue-almost every $\mathbf{c} \in (0, \bar{c})^K$,

$$G_{ij}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}; \mathbf{c}, b_{ij}) D_{\mathbf{x}_{ij} \mathbf{x}_{ij}}^2 N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \mathbf{c}_{ij}, b_{ij})$$

is invertible, and the bargaining solution is continuously differentiable w.r.t. $(\mathbf{x}_{-(ij)}, b_{ij})$ at that $\mathbf{x}_{-(ij)}$, with $b_{ij} \in (0, 1)$. This is true for all (ij) . Now we must extend the result for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$ for each (ij) , because the infinite union of Lebesgue-measure zero sets may not be a set of Lebesgue-measure zero.

Fix $(ij) \in G$. Let $\tilde{\mathbf{c}}_{ij}(\mathbf{x}_{-(ij)}) \subset (0, \bar{c})^K$ be the subset of Lebesgue-measure zero where continuous differentiability is not ensured by the application of transversality and implicit function theorems. Fix $\mathbf{x}_{-(ij)}$ and pick any $\mathbf{c} \in (0, \bar{c})^K \setminus \tilde{\mathbf{c}}_{ij}(\mathbf{x}_{-(ij)})$. By continuity of the bargaining solution, picking $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, for all $\varepsilon > 0$ there exists $\delta(\mathbf{x}_{-(ij)}, \varepsilon) > 0$ such that $\|\mathbf{x}'_{-(ij)} - \mathbf{x}_{-(ij)}\| < \delta(\mathbf{x}_{-(ij)}, \varepsilon)$ then $\|\mathbf{x}_{ij}^*(\mathbf{x}'_{-(ij)}; \mathbf{c}, b_{ij}) - \mathbf{x}_{ij}^*(\mathbf{x}_{-(ij)}; \mathbf{c}, b_{ij})\| < \varepsilon$. By continuity of the function that maps the entries of a matrix into its determinant, and because $N_{ij}(\cdot; b_{ij})$ is twice continuously differentiable in $\text{int}(B_{ij}(\mathbf{x}_{-(ij)}))$ (A2.1), there exists $\varepsilon(\mathbf{x}_{-(ij)}) > 0$ sufficiently small that for all $\mathbf{x}$ such that for all $\mathbf{x}'_{-(ij)} \in \mathbb{R}^{K-(ij)}$ such that $\|\mathbf{x}'_{-(ij)} - \mathbf{x}_{-(ij)}\| < \delta(\mathbf{x}_{-(ij)}, \varepsilon)$,

$$\det \left(D_{\mathbf{x}_{ij} \mathbf{x}_{ij}}^2 N_{ij}(\mathbf{x}_{ij}^*(\mathbf{x}'_{-(ij)}; \mathbf{c}_{ij}, b_{ij}), \mathbf{x}'_{-(ij)}; \mathbf{c}_{ij}, b_{ij}) \right) \neq 0.$$

Note that the open sets $Nb(\mathbf{x}_{-(ij)}, \varepsilon(\mathbf{x}_{-(ij)}), \delta(\mathbf{x}_{-(ij)}, \varepsilon))$ form an open cover of $X_{-(ij)}$. Since $X_{-(ij)}$ is compact, there exists a finite subopen cover, defined by $u = 1, \dots, U < \infty$ and the sets $Nb(\mathbf{x}_{-(ij)}^u, \varepsilon(\mathbf{x}_{-(ij)}^u), \delta(\mathbf{x}_{-(ij)}^u, \varepsilon))$. For each u the set $\tilde{\mathbf{c}}_{ij}(\mathbf{x}_{-(ij)})$ has Lebesgue-measure zero, so

$$\bigcup_{u=1}^U Nb(\mathbf{x}_{-(ij)}^u, \varepsilon(\mathbf{x}_{-(ij)}^u), \delta(\mathbf{x}_{-(ij)}^u, \varepsilon))$$

has Lebesgue-measure zero. So, for each $(ij) \in G$ the subset of $(0, \bar{c})^K \subset \mathbb{R}^K$ (and therefore of $[0, \bar{c}]^K \subset \mathbb{R}^K$) such that the bargaining solution is not continuously differentiable (and, also, such that Assumption 7 does not hold) has Lebesgue-measure zero. Taking the union of these sets over all $(ij) \in G$ yields a set of Lebesgue-measure zero too, because G is finite.

Since A7 holds, if A1 and 5 hold as well, we get twice continuous differentiability of the bargaining solution using the chain rule and the rule for the differentiation of matrix inverses, as in the proof of Proposition 5. ■

Proof of Lemma 5

Pick any $(ij) \in G$. By A2 Lemma 1 holds, and the bargaining solution is interior (A2.3) and the following FOC holds:

$$b_{ij} \frac{\partial \Delta \pi_{D_j}(\mathbf{x})}{\partial x_{ij}} \frac{1}{\Delta \pi_{D_j}(\mathbf{x})} + (1 - b_{ij}) \frac{\partial \Delta \pi_{U_i}(\mathbf{x})}{\partial x_{ij}} \frac{1}{\Delta \pi_{U_i}(\mathbf{x})} = 0. \quad (33)$$

By A2.2 both denominators are positive at the bargaining solution and, by assumption of Lemma 5, one of the partial derivatives is strictly increasing or decreasing. WLOG, let $\partial \Delta \pi_{D_j}(\mathbf{x}) / \partial x_{ij} < 0$, so

$\partial\Delta\pi_{U_i}(\mathbf{x})/\partial x_{ij} > 0$. Any increase in b_{ij} that is not compensated by a change in x_{ij} will make the left-hand side of (33) negative. Therefore, x_{ij}^* must have non-zero derivative with respect to b_{ij} . The same result holds if $\partial\Delta\pi_{D_j}(\mathbf{x})/\partial x_{ij} > 0$, exhausting all cases. ■

Proof of Lemma 6

Notation. Assumption 2 holds, so there exists at least one Nash-in-Nash equilibrium. Around each equilibrium, G_{ij} and V_{ij} , as defined in the proof of Proposition 6, are well defined. The diagonal $K \times K$ matrix $\tilde{G}$ stacks all the G_{ij} in its diagonal, and V stacks all the vector-valued V_{ij} , around each equilibrium. We only consider values of $\mathbf{c} \in (0, \bar{c})^K$ such that the bargaining solution functions are continuously differentiable w.r.t $\mathbf{x}_{-(ij)}$.

Strategy of proof. Again, we use the Poincaré-Hopf theorem.

Define $\tilde{Q} : X \times (0, \bar{c})^K \rightarrow \mathbb{R}^K$, a function of $\mathbf{x}$ parameterized by $\mathbf{c}$, as

$$\tilde{Q}(\mathbf{x}; \mathbf{c}) = \mathbf{x} - F(\mathbf{x}; \mathbf{c})$$

where we make the dependence on $\mathbf{c}$ explicit and leave dependence on $\mathbf{b}$ implicit. The proof that $\tilde{Q}$ points outwards at the boundary of X in the proof of Proposition 2 does not depend on $\mathbf{c}$, so it applies here to $\tilde{Q}$. It remains to prove that $D_{\mathbf{x}}\tilde{Q}(\mathbf{x}^*; \mathbf{c})$ is invertible at each Nash-in-Nash equilibrium $\mathbf{x}^* \in \text{int}(X)$, given $\mathbf{c}$.

Pick any $\mathbf{c}$ any equilibrium $\mathbf{x}^*$ yielded by such $\mathbf{c}$. From the proof of Proposition 6 we know that there exists an open subsets of X and $(0, \bar{c})^K$, $\tilde{X}$ and $\tilde{C}$, with $(\mathbf{x}^*, \mathbf{c}) \in \tilde{X} \times \tilde{C}$, such that for all $(\mathbf{x}, \mathbf{c}) \in \tilde{X} \times \tilde{C}$, is continuously differentiable w.r.t. $(\mathbf{x}, \mathbf{c})$. Note that

$$D_{\mathbf{x}}\tilde{Q}(\mathbf{x}; \mathbf{c}) = I - D_{\mathbf{x}}\tilde{G}^{-1}(\mathbf{x}; \mathbf{c}) \underbrace{V(\mathbf{x}; \mathbf{c})}_{=0} - \tilde{G}^{-1}(\mathbf{x}; \mathbf{c}) D_{\mathbf{x}}V(\mathbf{x}; \mathbf{c})$$

We use the transversality theorem. We want to prove that

$$[I - \tilde{G}^{-1}(\mathbf{x}; \mathbf{c}) D_{\mathbf{x}}V(\mathbf{x}; \mathbf{c}) | \tilde{G}^{-1}(\mathbf{x}; \mathbf{c}) D_{\mathbf{c}}V(\mathbf{x}; \mathbf{c})]$$

has full row rank. But $D_{\mathbf{c}}V(\mathbf{x}; \mathbf{c}) = I_K$ and $\tilde{G}^{-1}(\mathbf{x}; \mathbf{c})$ is square and invertible, so the full row rank condition holds. So, by the transversality theorem $D_{\mathbf{x}}\tilde{Q}(\mathbf{x}^*; \mathbf{c})$ is invertible for Lebesgue-almost every $\mathbf{c} \in \tilde{C}$.

The set of fixed points of F , E , is closed and, because it is a subset of X , E is bounded. So, it is compact. If the set of fixed points of F is finite, the proof is complete. Assume it is infinite. The union of open sets $\tilde{X}$ defined for each equilibrium above is an open cover of E . By compactness, there exists a finite subopen cover, defined by indices $t = 1, \dots, T < \infty$ and $\tilde{X}^t$. Each t has a corresponding $\tilde{C}^t$. Note that

$$\mathcal{C} = \bigcap_{t=1}^T \tilde{C}^t$$

is a finite intersection of open, nonempty sets, so $\mathcal{C}$ is open, nonempty, and contains $\mathbf{c}$. Let $\tilde{\mathcal{C}}^t$ the set of vectors $\mathbf{c} \in \tilde{C}$ such that, given $\mathbf{x}_{-(ij)}^*$, $D_{\mathbf{x}}\tilde{Q}(\mathbf{x}_{-(ij)}^*, \mathbf{c})$ is not invertible. That set has Lebesgue-measure

zero. Taking the union of such sets over t , we get a Lebesgue-measure zero set.

Thus, for Lebesgue-almost every element of $\mathcal{C}$, the invertibility condition necessary to apply the Poincaré-Hopf theorem holds. So, for each $\mathbf{c} \in (0, \bar{c})^K$, there exists an open subset $\mathcal{C} \subset (0, \bar{c})^K$ such that, for Lebesgue almost every $\mathbf{c}' \in \mathcal{C}$, if $\mathbf{c}'$ is the vector of suppliers' marginal costs, the model has a finite number of locally unique Nash-in-Nash equilibria. ■

Proof of Lemma 7

Pick any $(ij) \in G$. Because $K_{ij} = 1$, we reduce clutter by writing $\mathbf{x}$ as x and letting the U_i 's marginal cost be $c > 0$. Using the definition of q from the proof of Proposition 6, we can write $\Delta\pi_{U_i}(x, \mathbf{x}_{-(ij)}) = (x - c)q(x, \mathbf{x}_{-(ij)})$, with

$$q(x, \mathbf{x}_{-(ij)}) \equiv \sum_{l \in L_j} a_{hlj} s_l(\mathbf{x}) > 0$$

where the inequality holds because market shares are strictly positive in mixed logit (as long as there are consumers with remaining income).

With $b_{ij} = 1$ for all (ij) , because $\Delta\pi_{D_j}(\cdot, \mathbf{x}_{-(ij)})$ is strictly decreasing in x , we get that $x^*(\mathbf{x}_{-(ij)}; 1) = c$ for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$ for all (ij) , yielding a unique Nash-in-Nash equilibrium. For $b_{ij} \in (0, 1)$, by Lemma 1 the bargaining solution is interior and both firms have positive gains from trade. Taking the FOC of the bargaining problem in levels, and multiplying it by $\pi_{D_j}(x, \mathbf{x}_{-(ij)})^{1-b_{ij}} \pi_{U_i}(x, \mathbf{x}_{-(ij)})^{b_{ij}}$ we get an equivalent, nonsingularly-rescaled expression

$$R(x, \mathbf{x}_{-(ij)}; b_{ij}) = b_{ij} \pi_{U_i}(x, \mathbf{x}_{-(ij)}) \frac{\partial \pi_{U_i}(x, \mathbf{x}_{-(ij)})}{\partial x} + (1 - b_{ij}) \pi_{D_j}(x, \mathbf{x}_{-(ij)}) \frac{\partial \pi_{D_j}(x, \mathbf{x}_{-(ij)})}{\partial x} = 0 \quad (34)$$

Note that $R(c, \mathbf{x}_{-(ij)}; 1) = 0$ is well defined. By A2.1 R is continuously differentiable w.r.t. $(\mathbf{x}, b_{ij})$. Taking the derivative w.r.t. x we get

$$R_x(c, \mathbf{x}_{-(ij)}; 1) = \underbrace{\frac{\partial \pi_{U_i}(x, \mathbf{x}_{-(ij)})}{\partial x}}_{>0} \cdot \underbrace{\frac{\partial \pi_{D_j}(x, \mathbf{x}_{-(ij)})}{\partial x}}_{<0} < 0$$

because $\Delta\pi_{U_i}(c, \mathbf{x}_{-(ij)}) = 0$. Following the proof of Proposition 6, because $\pi_{D_j}(x, \mathbf{x}_{-(ij)})^{1-b_{ij}} \pi_{U_i}(x, \mathbf{x}_{-(ij)})^{b_{ij}} > 0$ (A2.2), the second derivative of the Nash product w.r.t. x is negative too. Using the implicit function theorem and (25) we get that, for each $(\mathbf{x}_{-(ij)}, b_{ij}) \in X_{-(ij)} \times (0, 1]$ in an open neighborhood around $(\mathbf{x}_{-(ij)}, \mathbf{b} = \mathbf{1})$ there is a unique, continuously differentiable bargaining solution. These neighborhoods generate an open cover of $X_{-(ij)}$. $X_{-(ij)}$ is compact and $\mathbf{b} = \mathbf{1}$ is fixed, so there exists a finite subopen cover. On a common neighborhood of that set, $|R_x|$ must be bounded away from zero.

By the application of the maximum theorem in the proof of Lemma 1, and because c is the unique bargaining solution if $b_{ij} = 1$, we get that $x^*(\mathbf{x}_{-(ij)}; b_{ij})$ is continuous on $X_{-(ij)} \times (0, 1]$. Let $\underline{b} \in (0, 1)$. Then, by the Heine-Cantor theorem $x^*(\mathbf{x}_{-(ij)}; b_{ij})$ is uniformly continuous on $X_{-(ij)} \times [\underline{b}, 1]$. Therefore, by the uniform convergence lemma in Cussen (2026), $x^*(\mathbf{x}_{-(ij)}; b_{ij})$ converges uniformly to $x^*(\mathbf{x}_{-(ij)}; 1)$ as $b_{ij} \uparrow 1$.

Since the bargaining solution is unique (A2, Lemma 1), as $b_{ij} \uparrow 1$ the bargaining solution matches

that yielded by using the implicit function theorem around $b_{ij} = 1$. Because

$$R(x, \mathbf{x}_{-(ij)}; b_{ij}) = D_x N_{ij}(x, \mathbf{x}_{-(ij)}; b_{ij}) \cdot \pi_{D_j}(x, \mathbf{x}_{-(ij)})^{1-b_{ij}} \pi_{U_i}(x, \mathbf{x}_{-(ij)})^{b_{ij}}$$

by the product rule we have that, around the bargaining solution,

$$D_{\mathbf{x}_{-(ij)}} R(x, \mathbf{x}_{-(ij)}; b_{ij}) = D_{x, \mathbf{x}_{-(ij)}}^2 N_{ij}(x, \mathbf{x}_{-(ij)}; b_{ij}) \pi_{D_j}(x, \mathbf{x}_{-(ij)})^{1-b_{ij}} \pi_{U_i}(x, \mathbf{x}_{-(ij)})^{b_{ij}}$$

But $D_{\mathbf{x}_{-(ij)}} \Delta \pi_{U_i}(c, \mathbf{x}_{-(ij)}, b_{ij}) = 0$ for all $\mathbf{x}_{-(ij)} \in X$, so $D_{x, \mathbf{x}_{-(ij)}}^2 N_{ij}(c, \mathbf{x}_{-(ij)}; b_{ij}) = 0$ for all $\mathbf{x}_{-(ij)} \in X$. Using (25) we have

$$D_{\mathbf{x}_{-(ij)}} x^*(\mathbf{x}_{-(ij)}; b_{ij}) = - \frac{D_{x, \mathbf{x}_{-(ij)}}^2 N_{ij}(x, \mathbf{x}_{-(ij)}; b_{ij})}{D_{x, x}^2 N_{ij}(x, \mathbf{x}_{-(ij)}; b_{ij})}$$

where the denominator is nonzero. Again, because $X_{-(ij)} \times [\underline{b}, 1]^{|G|}$ is compact, by the uniform convergence lemma from Cussen (2026) we have that

$$\|D_{\mathbf{x}_{-(ij)}} x^*(\mathbf{x}_{-(ij)}, b_{ij})\| \rightarrow 0 \quad \text{as} \quad b_{ij} \uparrow 1 \quad (35)$$

for all $(ij) \in G$. Because there are finitely many bargaining pairs, (35) implies there exists $\varepsilon \in (0, 1)$ such that if $\mathbf{b} \in (1 - \varepsilon, 1)^{|G|}$, every row sum of the absolute values of the off-diagonal entries of $D_{\mathbf{x}} F(\mathbf{x}; \mathbf{b}_{ij})$ is smaller than one. Then the conditions of Lemma 2 hold, Assumption 4 holds and, because A2 and A3.1 hold, by Proposition 3 the Nash in Nash equilibrium is globally unique. Finally, $(1 - \varepsilon, 1)^{|G|}$ has positive Lebesgue-measure. ■

Proof of Lemma 8

Consider a transformation of the bargaining problem. Assume D_j is a monopolist with RCNL demand that faces income constraint r_i for each product i it sells. Given the prices x_i of the 7 remaining stents, it must pick a price x_i greater than its marginal cost $c_i \geq 0$. Its objective function is

$$\tilde{\pi}(\mathbf{x}) = s_i(\mathbf{x})(x_i - c_i) + \sum_{h \neq i} s_h(\mathbf{x})(r_h - x_h) \quad \text{s.t.} \quad x_i \in [c_i, r_i].$$

The first partial derivative w.r.t. x_i is

$$\frac{\partial \tilde{\pi}}{\partial x_i} = s_i(\mathbf{x}) + (x_i - c_i) \frac{\partial s_i}{\partial x_i} + \sum_{h \neq i} \frac{\partial s_h}{\partial x_i} (r_h - x_h) \quad (36)$$

The price coefficient is constant and the nesting structure is simple (see Appendix D and the Online Appendix of Grennan (2013) cited there), so we can use the proofs of the derivatives of firm profits under random coefficients from Morrow and Skerlos (2010a). Ignoring the constraint $r_i \geq x_i$ for now, Corollary 2.18 of Morrow and Skerlos (2010a) states that, for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, there exists a finite $\bar{x}$, independent of $(x_h, r_i)_{h \in \mathcal{I} \setminus \{i\}}$ such that if $x_i > \bar{x}$, then (36) is negative. (Given that price sensitivity is not a random coefficient, for a proof that the conditions of Corollary 2.18 hold see Cussen (2026).) Define $\tilde{r}_i$ as such $\bar{x}$ for each (ij) .

It is clear that $\tilde{\pi}(r_i, \mathbf{x}_{-(ij)}) = \bar{\pi}_{D_j}(\mathbf{x}_{-(ij)})$. So if $\partial\tilde{\pi}/\partial x_i < 0$ for $x_i < r_i$ in a neighborhood of r_i , it must be that for such x_i $\tilde{\pi}(r_i, \mathbf{x}_{-(ij)}) > \bar{\pi}_{D_j}(\mathbf{x}_{-(ij)})$. But for such x_i it holds that

$$\begin{aligned}
0 + \bar{\pi}_{D_j}(\mathbf{x}_{-(ij)}) &= \bar{\pi}_{U_i}(\mathbf{x}_{-(ij)}) + \bar{\pi}_{D_j}(\mathbf{x}_{-(ij)}) \\
&= \tilde{\pi}(r_i, \mathbf{x}_{-(ij)}) \\
&< \tilde{\pi}(x_i, \mathbf{x}_{-(ij)}) \\
&= s_i(\mathbf{x})(x_i - c_i) + \sum_{h \neq i} s_h(\mathbf{x})(r_h - x_h) \\
&< \Delta\pi_{U_i}(x_i, \mathbf{x}_{-(ij)}) + \sum_{h \neq i} s_h(\mathbf{x})(r_h - x_h) + s_i(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)})(r_i - x_i) \\
&= \Delta\pi_{U_i}(x_i, \mathbf{x}_{-(ij)}) + \Delta\pi_{D_j}(x_i, \mathbf{x}_{-(ij)})
\end{aligned}$$

So, $\min\{\Delta\pi_{U_i}(x_i, \mathbf{x}_{-(ij)}), \Delta\pi_{D_j}(x_i, \mathbf{x}_{-(ij)})\} > 0$, the bargaining set is nonempty, and A2.2 holds for all (ij) for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$ if all r_i are sufficiently high. ■

B Gradient Descent Details

The code is written in Python, and we use JAX to compute derivatives. Our Python package contains the code details.²¹

B.1 Nash Product Computation

By Assumption 1 there exist functions $\mathbf{p}_j(\mathbf{x})$ that map first stage contracts into second stage strategies, such that a pair's Nash product is $\mathcal{N}_{ij}(\mathbf{x}, \mathbf{p}(\mathbf{x}); b_{ij})$. By Assumptions 1 and 5 these functions are twice continuously differentiable. We do not assume these functions admit closed form representation. Instead, we assume *first-order oracle access*: both the values of $\mathbf{p}_j(\mathbf{x})$ and $\mathcal{N}_{ij}(\mathbf{x}, \mathbf{p}(\mathbf{x}); b_{ij})$ and their gradients can be computed using, for example, automatic differentiation via JAX or PyTorch. This is a standard assumption.

B.2 Bargaining Solution Computation

To compute each dimension of $\tilde{F}_{ij}$ we need to solve an optimization problem over X_{ij} . At any candidate equilibrium $\mathbf{x}_t$, we maximize $\mathcal{N}_{ij}(\mathbf{x}_{ij}, \mathbf{x}; b_{ij})$ by fixing $\mathbf{x}_{-(ij)}$ and searching a grid representation of X_{ij} . This approach is slow if $\dim(X_{ij}) = K_{ij}$ is high. In typical applications, however, K_{ij} is small enough to not be a problem.²²

To compute $\mathcal{N}_{ij}$, we use its $\mathcal{N}_{it}$ representation. This is decomposed in two steps: (1) finding the equilibrium given $\mathbf{x}_{ij}$, and (2) computing $\mathcal{N}_{ij}$.

B.3 Equilibrium Computation

For logit, nested logit, mixed logit, and RCNL, this can be done quickly using the algorithm developed by Morrow and Skerlos (2011), implemented in `pyBLP` (Conlon and Gortmaker, 2020). For CES in monopolistic competition, one can use the constant markup property. In our package we compute Grennan (2013)'s equilibrium payoffs directly. Marginal cost functions used in empirical work (see (11)) are straightforward to compute and differentiate.

B.4 Gradient Computation

The key step is computing the Jacobian of M . Consider $\partial \tilde{F}_{ij} / \partial x_{ijk}$. If $(mn) = (ij)$ we have

$$\frac{\partial \tilde{F}_{ij}}{\partial x_{mnk}} = 0$$

²¹See code details at https://github.com/MoeenNehzati/nash_in_nash_GD. The total computation time is around 15 minutes in a laptop computer with 32 GB of RAM (5 GB used). CPU: 13th Gen Intel Core i7-1365U, 10 cores / 12 threads, 2 hardware threads per core, up to 5.2 GHz (lscpu).

²²See the papers cited in footnote 16. The paper with the largest K_{ij} we are able to find is Crawford and Yurukoglu (2012), where $K_{ij} \geq 6$ is possible. For other papers cited in Table 1, $K_{ij} > 1$ but small seems to be the norm. If the dimensions of the problem grow too large, it may be possible to rewrite a single bargaining problem as separate problems (say, because products are sold by separate divisions of a firm). This is similar to the approach in Grennan (2013) and, by reducing the dimensionality of each bargaining problem, can considerably speed up computation.

because $\tilde{F}_{ij}$ depends only on $\mathbf{x}_{-(ij)}$. For $(mn) \neq (ij)$, by the implicit function theorem (which by Proposition 5 we can use under Assumptions 2 and 7) we have, in matrix notation

$$\frac{\partial \tilde{F}_{ij}}{\partial \mathbf{x}_{mn}}(\mathbf{x}) = [D_{\mathbf{x}_{ij}, \mathbf{x}_{ij}}^2 N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; b_{ij})]^{-1} [D_{\mathbf{x}_{ij}, \mathbf{x}_{mn}}^2 N_{ij}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; b_{ij})] \quad (37)$$

By the chain rule we have, for all $(mn) \in G$, for all $k \in \mathcal{K}_{mn}$,

$$\begin{aligned} \frac{d\mathcal{N}_{ij}}{dx_{mnk}}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}, \mathbf{p}(\mathbf{x}); b_{ij}) &= \frac{\partial \mathcal{N}_{ij}}{\partial x_{mnk}}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}, \mathbf{p}(\mathbf{x}); b_{ij}) \\ &+ \sum_{r=1}^R \frac{\partial \mathcal{N}_{ij}}{\partial p_r}(\mathbf{x}_{ij}^*, \mathbf{x}_{-(ij)}, \mathbf{p}(\mathbf{x}); b_{ij}) \cdot \frac{\partial p_r(\mathbf{x})}{\partial x_{mnk}} \end{aligned} \quad (38)$$

where $R \equiv \sum_j \dim(P_j)$, and summing across $r \in \{1, \dots, R\}$ sums across all the elements of the strategy vectors of all retailers. A second application of the chain and product rules yield the second partial derivatives we need to construct (37).

B.5 Speed

The speed of computation of $\mathcal{N}_{ij}(\mathbf{x}, \mathbf{s}(\mathbf{x}); b_{ij})$ and the bargaining solution depends on the speed with which we can compute the second stage equilibrium. As mentioned above, researchers can use Morrow and Skerlos (2011)'s algorithm for discrete choice models. This algorithm is exceptionally fast and accurate.²³

Speed of computation of the gradient depends on how quickly one can differentiate. Our replication package shows in detail how to do this using automatic differentiation and Adam. The package implements our equilibrium computation for Grennan (2013)'s market for stents, but is designed to be adapted to other settings.

The speed of computation of $\tilde{F}_{ij}$ grows exponentially with $\max_{(ij) \in G} K_{ij}$. This improves over full grid search for X , which grows exponentially with K . This means that, holding the number of gridpoints per dimension constant, the computation cost using gradient descent can grow linearly in the number of bargaining pairs and of contract terms.

²³To make verification of A.2.3 easy, we also care about how fast one can code our algorithm. See Lee et al. (2021) and Gandhi and Nevo (2021) (Section 5.4) to see how our algorithm can recycle code used for estimation. Essentially, estimation of supplier marginal in bargaining models involves using GMM with the bargaining FOC as moments. So, if researchers use that estimation strategy expressions and code for that FOC will already exist.

C Full Grid Search Example

Let $I = J = 2$, $G = \mathcal{I} \times \mathcal{J}$. Each retailer is an intermediary and sells two products, buying one from each supplier. So, $K = 4$, $K_{ij} = 1$ for all (ij) and we can set the lower bound of each X_{ij} to zero without loss. We impose a maximum $\bar{c}_{ij} > 0$ in each dimension. With finite resources this is without loss. It can also be interpreted as the cost for retailers of buying an input from an alternative, inefficient seller (Hristakeva, 2022), or producing in-house. So, $X_{ij} = [0, \bar{c}_{ij}]$.

Suppliers' costs are normalized to zero, and retailers' costs are linear in input prices and quantities. The demand system is mixed logit with population normalized to one. On and off-path profits are

$$\begin{aligned}\hat{\pi}_{D_j}(x_{ij}, \mathbf{x}_{-(ij)}) &= \sum_{i=1,2} (p_{ij} - x_{ij}) \cdot s_{ij}(\mathbf{p}) \\ \hat{\pi}_{D_j}(\mathbf{x}_{-(ij)}) &= (p_{ij} - \bar{c}_{ij}) \cdot s_{ij}(\mathbf{p}^{\text{off}}) + (p_{-ij} - x_{-ij}) \cdot s_{-ij}(\mathbf{p}^{\text{off}}) \\ \hat{\pi}_{U_i}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}) &= \sum_{j=1,2} x_{ij} \cdot s_{ij}(\mathbf{p}) \\ \hat{\pi}_{U_i}(\mathbf{x}_{-(ij)}) &= x_{i-j} \cdot s_{i-j}(\mathbf{p}^{\text{off}}).\end{aligned}$$

where $\mathbf{p}^{\text{off}}$ is the equilibrium price in the off-path case where bargaining for pair (ij) breaks down, and s_{ij} is the market share of the product D_j buys from U_i .

Figure 1 shows a representative plot of the Nash product function for pair (11) given some $\mathbf{x}_{-(ij)}$. The single-peakedness within the grid (indeed, the hump-shape we see in simple Nash product functions) is evident, and we verify it for each (ij) for each $\mathbf{x}_{-(ij)}$ in the grid. Alternative cost functions, distributions of price sensitivities, and grids, do not affect the result.

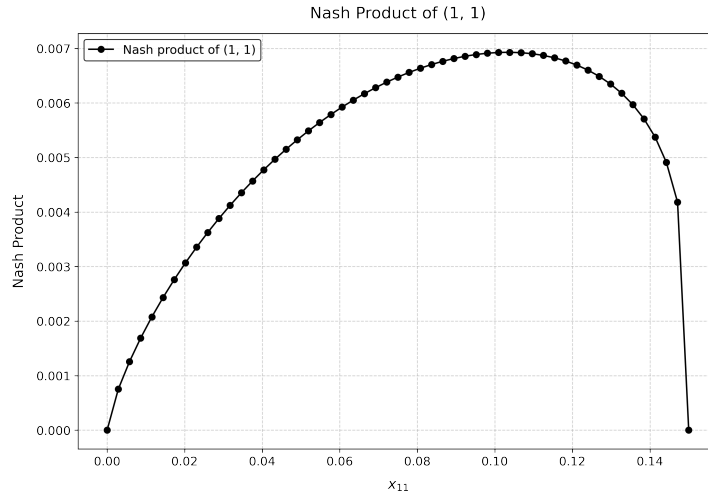


Figure 1: Representative $N_{11}(x_{ij}, \mathbf{x}_{-(11)}; b_{11})$ with 7,890,481 gridpoints.

The code can be downloaded at <https://github.com/diegocussen/bargaining-equilibrium-exist>. Computation of the largest model (7,890,481 gridpoints) takes approximately six hours in a MacBook Air with 8 GB of RAM and an Apple M2 chip.

D Computational Exercise Details

In this appendix we use Grennan (2013) and its Online Appendix.²⁴ Let j be the hospital, i the stent. There is a single hospital, so we save in notation by dropping the j when convenient. Since there is the most data for September of 2005 (Tables 1 and 5 of that paper), we will use that month as our reference. There are 6 *BMS* and 2 *DES* stents in that month, so $|G| = 8$. For each stent i , mean utilities are

$$\delta_i(x_i) = \theta_i - \theta_p x_i + \xi_i, \quad \delta_0 = 0.$$

Each doctor-patient pair is loyal to one stent i , reflecting the doctor's belief that a specific stent is particularly well suited to a specific patient. There are as many loyalty types L as stents: $L = I$. Following the Appendix, we set the probability of a patient being loyal to a specific stent i , ϕ_i , as the market share of that stent, conditional on getting a stent. We do not have empirical market shares, so we roughly back them out from Figure 1, panel A for *BMS* and *DES*. Within each type of stent, we divide the share equally. These numbers are, roughly, $5000/5500 = 0.9091$ for *BMS* and $1 - 0.9091$ for *DES*. We get

$$\phi^l = \begin{cases} 0.9091/6 & \text{if } l \text{ is } bms \\ (1 - 0.9091)/2 & \text{if } l \text{ is } des. \end{cases}$$

As the Appendix points out, if $\lambda = 0$ or $\lambda \gg 0$, then the distribution (i.e., the ϕ^l) matters little. This is the case here, so our assumptions on ϕ are not critical.

We construct market share functions using the nested logit equations in the Appendix. To parameterize them, we use the estimated parameter values reported in Grennan (2013) and its Appendix. The data are proprietary, so as to not replicate prices, the θ_i are not reported. Instead, we draw the θ_i iid from a uniform distribution.

We use average estimated marginal costs reported in the paper as our measure of supplier marginal cost. We do not draw from a distribution because it is right-skewed with unknown skewness and a non-negativity constraint. So, marginal costs are $c_i = 34$ for *BMS* and $c_i = 1,103$ for *DES*.

It remains to provide a measure of revenue per angioplasty. Revenue is independent of cost, but markup is not. To construct the revenue, consider the revenue the hospital gets from a stent procedure for patient n . Grennan (2013) presents evidence that r_n depends on n 's insurance (or being part of Medicare), but not on the type of stent (beyond whether it is *BMS* or *DES*). Using invariance to positive affine transformations (Lemma 1), we assume there is a continuum of mass one of patients and write D_j 's on path profits as

$$\hat{\pi}_{D_j}(\mathbf{x}) = \sum_{i:(ij) \in G} \left(\int_{n \in A_i} r_n dn - x_i \right) s_i(\mathbf{x})$$

where $n \in A_i$ means that stent i is chosen for patient n . So, the first term of the integral yields average revenue. In line with the error structure of the model, we assume the best stent for a patient is orthogonal to the revenue they bring (equation 4. p. 156). Since all we need is average revenue, all we need is Medicare revenue (provided by Grennan (2013)), the fraction of patients who are in Medicare,

²⁴Online appendix for that paper can be found at <https://www.aeaweb.org/articles?id=10.1257/aer.103.1.145>.

and the average markup private insurance firms offer as reimbursement over Medicare. From various sources we get the following:

- Medicare revenue for *BMS*:

$$r_n = \underbrace{812}_{\text{reimbursement for doctor}} + \underbrace{10,422}_{\text{reimbursement for hospital}}$$

- Medicare revenue for *DES*:

$$r_n = 812 + \underbrace{11,814}_{\text{reimbursement for hospital}}$$

- Fraction of US population in Medicare. We use September 2005 because we have the most numbers for that year/month in the paper. US population in 2005: 295,516,599.²⁵ Number of Americans in Medicare in 2005: 42.6 million.²⁶ So, fraction of US population in Medicare = 14.41%.
- Average markup of private insurance over Medicare. We did not find numbers from 2005, so we use the Congressional Budget Office's (CBO) 2018 report, plus numbers from Grennan (2013), to get an average markup of approximately 46%.

We get

$$\bar{r}_i = \int_{n \in A_i} r_n dn = \begin{cases} 15,656.98 & \text{if } i \text{ is } bms \\ 17,597.03 & \text{if } i \text{ is } des \end{cases}$$

For this computation, we have $r_n^{bms} = (812 + 10,422) \cdot (0.1441 + 1.46[1 - 0.1441])$ and $r_n^{des} = (812 + 11,814) \cdot (0.1441 + 1.46[1 - 0.1441])$.

²⁵<https://www.macrotrends.net/global-metrics/countries/usa/united-states/population>. Last accessed on October 4, 2025.

²⁶US Department of Health and Human Services, 2007 CMS Statistics. Retrieved at https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMS-Statistics-Reference-Booklet/Downloads/2007CMSStatistics.pdf&ved=2ahUKEwjs1JHmpZCQAxWeMVkFHavLHAQQFnoECBgQAQ&usg=AOvVaw0x5m-U4f_MpYyDJ_XODk8I. Last accessed on October 6, 2025.

E Continuous Differentiability with Discrete Choice Demand

By discrete choice demand we mean logit, nested logit, mixed logit or RCNL. Rewriting (10) and (12) we have

$$\pi_{D_j}(\mathbf{p}_j, \mathbf{p}_{-j}; \mathbf{x}, \boldsymbol{\theta}) = \sum_{l \in L_j} (p_l - mc_l(\mathbf{x}), \boldsymbol{\theta}) s_l(\mathbf{p}, \boldsymbol{\theta})$$

$$\pi_{U_i}(\mathbf{p}, \mathbf{x}) = \sum_{j: (ij) \in G} \sum_{l \in L_j} \sum_{h \in H_i} a_{jlh} s_l(\mathbf{p}; \boldsymbol{\theta}) (x_h - c_h(\boldsymbol{\theta})).$$

Let $\boldsymbol{\theta}$ contain suppliers' marginal costs c_h , the parameters γ , ω , and ν in retailers' marginal cost functions (11), and parameters in consumers' mean utilities (considering linear utilities). These parameters are sufficient to achieve the proper parameterization required in Appendix F.

Because market shares are continuous in prices, and using either the identity or log version of (11), all profit functions are infinitely continuously differentiable w.r.t. all c_h , γ , ω , and ν . Market shares are infinitely continuously differentiable w.r.t the parameters in mean utilities (Iaria and Wang, 2019; Conlon and Gortmaker, 2020). Profits are continuous functions of market shares, yielding continuous differentiability of profits.

For linear input prices and two-part tariffs the same argument that shows infinite continuous differentiability of profits w.r.t. γ , ω , and ν proves infinite continuous differentiability w.r.t. $\mathbf{x}$.

Finally, regarding $\mathbf{p}$, the proof of n th-order continuous differentiability of profits w.r.t. all prices under mixed logit with linear demand in Cussen (2026) is directly established for logit in that paper too, and applies to nested logit as well.

F Equivalent Bargaining-Set Formulations

Throughout, fix a bargaining pair (ij) and $\mathbf{x}_{-(ij)} \in X_{-(ij)}$. The argument has three parts. First, we provide an abstract condition, stated in terms of the bargaining pair's gains from trade, under which the three bargaining set formulations coincide for almost every primitive parameter. This condition is potentially applicable across a broad range of contracting environments because it is stated solely in terms of the bargaining pair's gains from trade.

This generality, however, leaves open whether the condition is satisfied in any particular model; that must be verified from the model's primitives. Second, we perform this verification in a logit model using relationship-local, product-level demand or cost shifters together with a supplier-cost shifter. Third, we use parameter values at which mixed logit reduces to logit. Under reasonable regularity conditions, we show that the rank condition extends locally around any compatible logit anchor.

Note that our proofs are conditional on $\mathbf{x}_{-(ij)} \in X_{-(ij)}$. To see how they extend to all $X_{-(ij)}$ (for all $(ij) \in G$), let $S(\mathbf{x}_{-(ij)})$ be the subset of Θ such that the bargaining set may have empty interior if $\theta \in S(\mathbf{x}_{-(ij)})$. Because the bargaining set correspondence is continuous in $\mathbf{x}_{-(ij)}$ (proof of Lemma 1), for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$ if $\text{int}(B_{ij}(\mathbf{x}_{-(ij)}; \theta)) \neq \emptyset$, then there exists a ball in $\mathbb{R}^{K-(ij)}$ of radius $\varepsilon_{\mathbf{x}_{-(ij)}}$, denoted $b(\mathbf{x}_{-(ij)}, \varepsilon_{\mathbf{x}_{-(ij)}})$ such that, for all $\mathbf{x}'_{-(ij)} \in b(\mathbf{x}_{-(ij)}, \varepsilon_{\mathbf{x}_{-(ij)}})$, $\text{int}(B_{ij}(\mathbf{x}'_{-(ij)}; \theta)) \neq \emptyset$. Thus, for each $\mathbf{x}_{-(ij)}$, the subset of Θ such that, for any $\mathbf{x}'_{-(ij)} \in b(\mathbf{x}_{-(ij)}, \varepsilon_{\mathbf{x}_{-(ij)}})$, $\text{int}(B_{ij}(\mathbf{x}'_{-(ij)}; \theta))$ may be empty, is a subset of $S(\mathbf{x}_{-(ij)})$, and thus has Lebesgue measure zero.

$X_{-(ij)}$ is compact and the $b(\mathbf{x}_{-(ij)}, \varepsilon_{\mathbf{x}_{-(ij)}})$ form an open cover of it, so there exists a finite subopen cover, indexed by $u = 1, \dots, U < \infty$, and composed of $b(\mathbf{x}_{-(ij)}^u, \varepsilon_{\mathbf{x}_{-(ij)}^u})$. Then, if $\text{int}(B_{ij}(\mathbf{x}'_{-(ij)}; \theta)) = \emptyset$, it must be that

$$\theta \in S = \bigcup_{u=1}^U S(\mathbf{x}_{-(ij)}^u)$$

which is a finite union of Lebesgue-measure zero sets. So, S has Lebesgue-measure zero. This is true for all (ij) . Therefore, if our conditions stated below hold, the set of parameters such that Assumption 2.2 does not hold in a model has Lebesgue-measure zero.

F.1 Three bargaining-set formulations and abstract equivalence

Let G be the contracting graph containing (ij) , and write $G_{-(ij)} := G \setminus \{(ij)\}$. The two relevant graphs are G and $G_{-(ij)}$. Fix the contracts $\mathbf{x}_{-(ij)}$ on all other relationships. In the paper's notation, the directed gains from adding (ij) are

$$\Delta\pi_{D_j}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \theta) := \hat{\pi}_{D_j}(\mathbf{x}, \theta) - \bar{\pi}_{D_j}(\mathbf{x}_{-(ij)}, \theta), \quad \Delta\pi_{U_i}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \theta) := \hat{\pi}_{U_i}(\mathbf{x}, \theta) - \bar{\pi}_{U_i}(\mathbf{x}_{-(ij)}, \theta).$$

Maintain the paper's contract space $X_{ij} \subset \mathbb{R}^{K_{ij}}$. For the transversality argument, suppose in addition that

$$X_{ij} = \{\mathbf{x}_{ij} : g_m(\mathbf{x}_{ij}) \geq 0, \ m = 1, \dots, M_{ij}\},$$

where $M_{ij} \in \mathbb{N}$, each g_m is defined and $C^{K_{ij}}$ on a neighborhood of X_{ij} , and the gradients of the binding constraints are linearly independent at every point of X_{ij} (LICQ). This condition holds if X_{ij} is a box in $\mathbb{R}^{K_{ij}}$ which is the case for all the contract types in Lemma 3, among others. Let $\Theta^\circ \subset \text{int } \Theta$ be

open, and assume that the gains are $C^{K_{ij}}$ on a common open neighborhood of $X_{ij} \times \Theta^\circ$. For the fixed conditioning profile $\mathbf{x}_{-(ij)}$, define the stacked gain map

$$\Delta\pi_{ij, \mathbf{x}_{-(ij)}} : X_{ij} \times \Theta^\circ \longrightarrow \mathbb{R}^2, \quad \Delta\pi_{ij, \mathbf{x}_{-(ij)}}(\mathbf{x}_{ij}, \boldsymbol{\theta}) := (\Delta\pi_{D_j}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}), \Delta\pi_{U_i}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta})).$$

We retain the notation $\Delta\pi_{ij, \mathbf{x}_{-(ij)}}$ when specializing the parameter domain below. Define

$$\begin{aligned} B_{ij}^1(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}) &:= cl_{ij}\{\mathbf{x}_{ij} : \Delta\pi_{D_j}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) > 0, \Delta\pi_{U_i}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) > 0\}, \\ B_{ij}^2(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}) &:= cl_{ij}\{\mathbf{x}_{ij} : \Delta\pi_{D_j}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) > 0\} \cap cl_{ij}\{\mathbf{x}_{ij} : \Delta\pi_{U_i}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) > 0\}, \\ B_{ij}^3(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}) &:= \{\mathbf{x}_{ij} \in X_{ij} : \Delta\pi_{D_j}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) \geq 0, \Delta\pi_{U_i}(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) \geq 0\}. \end{aligned}$$

Thus B_{ij}^1 is the joint-strict-closure formulation, $B_{ij}^2(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}) = B_{ij}(\mathbf{x}_{-(ij)}; \boldsymbol{\theta})$ is the paper's separate-closure formulation from (1), and $B_{ij}^3(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}) = \tilde{B}_{ij}^\circ$ is the weak-gain formulation introduced in the main text. Continuity always gives

$$B_{ij}^1(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}) \subseteq B_{ij}^2(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}) \subseteq B_{ij}^3(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}).$$

These inclusions can be strict even in the one-dimensional case $X_{ij} = [-1, 1]$. If $\Delta\pi_{ij, \mathbf{x}_{-(ij)}}(x) = (x, -x)$, then $B_{ij}^1 = \emptyset$ while $B_{ij}^2 = B_{ij}^3 = \{0\}$. If $\Delta\pi_{ij, \mathbf{x}_{-(ij)}}(x) = (-x^2, 1)$, then $B_{ij}^1 = B_{ij}^2 = \emptyset$ while $B_{ij}^3 = \{0\}$. Thus the only issue is whether every weakly profitable boundary point can be approximated by contracts at which both gains are strictly positive.

Definition F.1 (Proper parameterization). The gain family is *properly parameterized* for the fixed pair and conditioning profile $((ij), \mathbf{x}_{-(ij)})$ if

$$D_{\boldsymbol{\theta}} \Delta\pi_{ij, \mathbf{x}_{-(ij)}}(\mathbf{x}_{ij}, \boldsymbol{\theta}) : \mathbb{R}^{\dim(\Theta)} \longrightarrow \mathbb{R}^2 \quad \text{is onto for every } (\mathbf{x}_{ij}, \boldsymbol{\theta}) \in X_{ij} \times \Theta^\circ. \quad (\text{PR})$$

Thus parameter variation alone gives the pair of directed gains full rank.

We next connect this submersion condition to the transversality used in the proof. Let $\mathcal{S}_{ij}$ be the finite collection of exact active-set strata of X_{ij} : for every realized active set $A \subseteq \{1, \dots, M_{ij}\}$, let

$$S_A := \{\mathbf{x}_{ij} \in X_{ij} : g_m(\mathbf{x}_{ij}) = 0 \iff m \in A\}, \quad \mathcal{S}_{ij} := \{S_A : S_A \neq \emptyset\}.$$

Every $\mathbf{x}_{ij} \in X_{ij}$ lies in a unique $S \in \mathcal{S}_{ij}$, and LICQ makes each S a smooth manifold; write $T_{\mathbf{x}_{ij}}S$ for its tangent space at $\mathbf{x}_{ij}$. Restricting the contract domain to S removes only contract directions and leaves all parameter directions available. Hence (PR) implies that the restrictions of $\Delta\pi_{ij, \mathbf{x}_{-(ij)}}$ and both of its components to $S \times \Theta^\circ$ remain submersions. In particular,

$$\Delta\pi_{D_j} \Big|_{S \times \Theta^\circ} \pitchfork \{0\}, \quad \Delta\pi_{U_i} \Big|_{S \times \Theta^\circ} \pitchfork \{0\}, \quad \Delta\pi_{ij, \mathbf{x}_{-(ij)}} \Big|_{S \times \Theta^\circ} \pitchfork \{(0, 0)\}.$$

Here $\pitchfork$ denotes transversality to the indicated zero set. Parametric transversality will transfer these properties to their contract sections for almost every fixed $\boldsymbol{\theta}$.

Lemma F.1 (Deterministic equivalence). Fix θ . Suppose that, for every $S \in \mathcal{S}_{ij}$,

$$\Delta\pi_{D_j}(\cdot, \mathbf{x}_{-(ij)}; \theta)|_S \not\supset \{0\}, \quad \Delta\pi_{U_i}(\cdot, \mathbf{x}_{-(ij)}; \theta)|_S \not\supset \{0\}, \quad \Delta\pi_{ij, \mathbf{x}_{-(ij)}}(\cdot, \theta)|_S \not\supset \{(0, 0)\}.$$

Then

$$B_{ij}^1(\mathbf{x}_{-(ij)}; \theta) = B_{ij}^2(\mathbf{x}_{-(ij)}; \theta) = B_{ij}^3(\mathbf{x}_{-(ij)}; \theta).$$

Proof. It is enough to prove $B_{ij}^3(\mathbf{x}_{-(ij)}; \theta) \subseteq B_{ij}^1(\mathbf{x}_{-(ij)}; \theta)$. Fix $\mathbf{x}_{ij} \in B_{ij}^3(\mathbf{x}_{-(ij)}; \theta)$, and let $S \in \mathcal{S}_{ij}$ be the unique stratum containing $\mathbf{x}_{ij}$. If both gains are positive, there is nothing to prove.

If one gain is zero and the other is positive, transversality of the zero gain on S gives a direction $d \in T_{\mathbf{x}_{ij}}S$ along which its derivative is positive. Along a C^1 curve in S tangent to d , the zero gain becomes positive while the other remains positive.

If both gains are zero, transversality of $\Delta\pi_{ij, \mathbf{x}_{-(ij)}}$ means

$$D_{\mathbf{x}_{ij}}\Delta\pi_{ij, \mathbf{x}_{-(ij)}}(\mathbf{x}_{ij}, \theta)|_{T_{\mathbf{x}_{ij}}S} : T_{\mathbf{x}_{ij}}S \longrightarrow \mathbb{R}^2$$

is onto. Hence some direction $d \in T_{\mathbf{x}_{ij}}S$ is mapped to $(1, 1)$. A curve in S tangent to d makes both gains positive. In every case, $\mathbf{x}_{ij}$ is a limit of points in the joint strict-positivity set and therefore belongs to $B_{ij}^1(\mathbf{x}_{-(ij)}; \theta)$. ■ □

Theorem F.1 (Almost-everywhere equivalence). If the gain family is properly parameterized, then

$$B_{ij}^1(\mathbf{x}_{-(ij)}; \theta) = B_{ij}^2(\mathbf{x}_{-(ij)}; \theta) = B_{ij}^3(\mathbf{x}_{-(ij)}; \theta)$$

for Lebesgue-almost every $\theta \in \Theta^\circ$.

Proof of Theorem F.1. Proper parameterization implies the three lifted transversality conditions displayed above. The finite-dimensional parametric transversality theorem (Hirsch, 2012) states that if a C^q map $\Psi : \mathcal{M} \times \Theta^\circ \rightarrow \mathbb{R}^k$ is transverse to zero and $q > \max\{0, \dim \mathcal{M} - k\}$, then $\Psi(\cdot, \theta)$ is transverse to zero for Lebesgue-almost every θ . Apply it to each of the three restricted maps on every $S \in \mathcal{S}_{ij}$. Since $\dim S \leq K_{ij}$ and $K_{ij} > 0$, regularity of class $C^{K_{ij}}$ is sufficient for every scalar and stacked map. There are finitely many active-set strata, so one may intersect the resulting full-measure parameter sets. Thus the hypotheses of Lemma F.1 hold for almost every θ , which proves the result. ■

Remark F.1 (Scope). At a fixed contract profile $\mathbf{x}$, and hence fixed $\mathbf{x}_{-(ij)}$ for each bargaining pair (ij) , suppose that every pair's gain family is properly parameterized on a common parameter domain. Finiteness of the graph then implies that, for almost every θ , the three definitions coincide for all bargaining pairs simultaneously. This finite-intersection argument need not extend to a continuum of pairs.

This completes the general argument. Next we verify its gain-rank condition in logit and mixed logit.

F.2 Verification in Logit Models

This section keeps the bargaining-set definitions fixed and adds multiproduct Bertrand competition under logit demand. Its only goal is to verify proper parameterization.

F.2.1 Gains from logit primitives

Use $G' \in \{G, G_{-(ij)}\}$ as the graph-state index and define its contract profile by

$$\mathbf{x}^G := \mathbf{x} = (\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}), \quad \mathbf{x}^{G_{-(ij)}} := \mathbf{x}_{-(ij)}.$$

For each G' , let $\mathbf{v}^{G'}(\mathbf{x}^{G'}, \boldsymbol{\theta})$ denote the vector of the price-independent parts of products' mean utilities, $\mathbf{c}^{G'}(\mathbf{x}^{G'}, \boldsymbol{\theta})$ the vector of product marginal costs, and $\mathbf{s}^{G'}(\mathbf{p}; \mathbf{x}^{G'}, \boldsymbol{\theta})$ the vector of demand shares. Retailers then play the second stage multiproduct Bertrand pricing game; $\mathbf{p}^{*,G'}(\mathbf{x}^{G'}, \boldsymbol{\theta})$ denotes its selected equilibrium. Consistently with the paper's payoff notation, write $\pi_f^{G'}(\mathbf{p}, \mathbf{x}^{G'}; \boldsymbol{\theta})$ for firm f 's payoff at arbitrary prices, and define the equilibrium payoffs by

$$\begin{aligned} \hat{\pi}_f(\mathbf{x}_{ij}, \mathbf{x}_{-(ij)}; \boldsymbol{\theta}) &:= \pi_f^G(\mathbf{p}^{*,G}(\mathbf{x}, \boldsymbol{\theta}), \mathbf{x}; \boldsymbol{\theta}), \\ \bar{\pi}_f(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}) &:= \pi_f^{G_{-(ij)}}(\mathbf{p}^{*,G_{-(ij)}}(\mathbf{x}_{-(ij)}, \boldsymbol{\theta}), \mathbf{x}_{-(ij)}; \boldsymbol{\theta}). \end{aligned}$$

These objects determine $\Delta\pi_{ij, \mathbf{x}_{-(ij)}}(\mathbf{x}_{ij}, \boldsymbol{\theta})$ through the gain definition above. Throughout this section, we omit the displayed arguments of these objects when no ambiguity arises; their dependence on contracts and parameters remains in force. Logit demand specifies how mean utilities and prices determine shares, but not how contracts affect mean utilities, marginal costs, or payoffs. That link must be part of the primitive family. We give one relationship-local affine specification for which $\Delta\pi_{ij, \mathbf{x}_{-(ij)}}$ satisfies (PR).

Let $\mathcal{J}$ be the finite set of retailers. Recall that $j \in \mathcal{J}$ is the retailer in pair (ij) . Each $r \in \mathcal{J}$ owns a nonempty product set L_r , and $(L_r)_{r \in \mathcal{J}}$ is a disjoint partition of $\mathcal{L}$, where $L = |\mathcal{L}|$. Homogeneous-logit demand shares under G' are

$$s_l^{G'} = \frac{\exp(v_l^{G'} - \alpha p_l)}{1 + \sum_{k \in \mathcal{L}} \exp(v_k^{G'} - \alpha p_k)}, \quad \alpha > 0,$$

where the outside utility is normalized to zero, and

$$s_0^{G'} := \frac{1}{1 + \sum_{k \in \mathcal{L}} \exp(v_k^{G'} - \alpha p_k)}, \quad \sigma_r^{G'} := \sum_{l \in L_r} s_l^{G'}, \quad r \in \mathcal{J}.$$

Here $s_0^{G'}$ is the outside-option share, while $\sigma_r^{G'}$ is retailer r 's aggregate share; in particular, $\sigma_j^{G'}$ is the aggregate share of the retailer in pair (ij) . For the finite product set considered here, finite utilities and prices imply $s_l^{G'} > 0$ and $s_0^{G'} > 0$ directly from these formulas. The constant α is the homogeneous-logit specialization of the paper's $\alpha(\theta)$. Restrict attention to an open parameter domain $\Theta^{HL} \subseteq \Theta^\circ$ on which, for every $\mathbf{x}_{ij} \in X_{ij}$ and $G' \in \{G, G_{-(ij)}\}$, the selected multiproduct-Bertrand equilibrium is an interior solution of the unconstrained price first-order conditions and is jointly $C^{K_{ij}}$ in $(\mathbf{x}_{ij}, \boldsymbol{\theta}^{HL})$. These are the interiority and regularity conditions used below.

Each retailer's downstream pricing objective is its standard variable profit with constant per-unit product costs. Normalize market size to one, equivalently measuring all payoffs and payoff parameters per unit of market size. For retailer D_j , define variable profit by

$$\pi_{D_j, \text{var}}^{G'} := \sum_{l \in L_j} (p_l - c_l^{G'}) s_l^{G'},$$

and, for $l \in L_j$, the product-level first-order condition and logit share derivative are

$$\begin{aligned} 0 &= s_l^{G'} + \sum_{k \in L_j} (p_k - c_k^{G'}) \frac{\partial s_k^{G'}}{\partial p_l}, \\ \frac{\partial s_k^{G'}}{\partial p_l} &= -\alpha s_k^{G'} (\mathbf{1}[k = l] - s_l^{G'}). \end{aligned}$$

Dividing the first equation by $s_l^{G'} > 0$ gives

$$p_l - c_l^{G'} = \frac{1}{\alpha} + \sum_{k \in L_j} (p_k - c_k^{G'}) s_k^{G'}.$$

The right-hand side is independent of l , so all products of retailer D_j have the common equilibrium markup

$$\frac{1}{\alpha(1 - \sigma_j^{G'})}.$$

Thus

$$\pi_{D_j, \text{var}}^{G'} = V(\sigma_j^{G'}), \quad V(\sigma) := \frac{1}{\alpha} \frac{\sigma}{1 - \sigma}, \quad \sigma \in (0, 1), \quad V'(\sigma) = \frac{1}{\alpha(1 - \sigma)^2} > 0,$$

where prices and shares are evaluated at $\mathbf{p}^{*, G'}$ whenever their arguments are suppressed. Any additional terms in the retailer's bargaining payoff are assumed to be price-independent and identical under G and $G_{-(ij)}$; hence they cancel from the gain. Therefore

$$\Delta \pi_{D_j} = V(\sigma_j^G) - V(\sigma_j^{G_{-(ij)}}).$$

F.2.2 Verification under relationship-local affine parameterization

Continue to fix $(ij) = (ij)$ and $\mathbf{x}_{-(ij)}$ as above, and let $G' \in \{G, G_{-(ij)}\}$. The contract $\mathbf{x}_{ij}$ ranges over X_{ij} . Assume that product availability and ownership remain fixed across the two graph states, and let the nonempty set $L_{ij} \subseteq L_j$ collect the products whose primitives may be shifted by pair (ij) .

Within the homogeneous-logit submodel, write

$$\boldsymbol{\theta}^{HL} = (\boldsymbol{\vartheta}^{HL}, \lambda_{ij}, \rho_{ij}, \kappa_{ij}) \in \Theta^{HL},$$

where the superscript HL labels the homogeneous-logit submodel. The block $\boldsymbol{\vartheta}^{HL}$ collects all remaining parameters; it may enter the primitives nonlinearly. The blocks $(\lambda_{ij}, \rho_{ij}, \kappa_{ij})$ contain independent

intercept shifters for (ij) . For products owned by D_j , write

$$\lambda_{ij} = (\lambda_{l,ij})_{l \in L_{ij}} \in \mathbb{R}^{|L_{ij}|}, \quad \rho_{ij} = (\rho_{l,ij})_{l \in L_{ij}} \in \mathbb{R}^{|L_{ij}|}, \quad \kappa_{ij} \in \mathbb{R},$$

where $\lambda_{l,ij}$ and $\rho_{l,ij}$ are product-specific utility and marginal-cost shifters.

For each product $l \in L_{ij}$, suppose

$$v_l^{G'}(\mathbf{x}^{G'}, \boldsymbol{\theta}^{HL}) = v_l^{0,G'}(\mathbf{x}^{G'}, \boldsymbol{\vartheta}^{HL}) + \mathbf{1}[(ij) \in G'] \lambda_{l,ij}, \quad (40)$$

$$c_l^{G'}(\mathbf{x}^{G'}, \boldsymbol{\theta}^{HL}) = c_l^{0,G'}(\mathbf{x}^{G'}, \boldsymbol{\vartheta}^{HL}) + \mathbf{1}[(ij) \in G'] \rho_{l,ij}. \quad (41)$$

For products outside L_{ij} , the displayed pair- (ij) coefficients have no effect. The baseline primitives may depend arbitrarily on $\boldsymbol{\vartheta}^{HL}$ and on the relevant contracts, subject to the standing smoothness conditions, but are independent of $(\lambda_{ij}, \rho_{ij}, \kappa_{ij})$.

The supplier-payoff shifter for pair (ij) is κ_{ij} , measured in the same per-market payoff units as profits. At arbitrary prices, its payoff satisfies

$$\pi_{U_i}^{G'}(\mathbf{p}, \mathbf{x}^{G'}; \boldsymbol{\theta}^{HL}) = \pi_{U_i}^{0,G'}(\mathbf{p}, \mathbf{x}^{G'}; \boldsymbol{\vartheta}^{HL}, \lambda_{ij}, \rho_{ij}) - \mathbf{1}[(ij) \in G'] \kappa_{ij}, \quad (42)$$

where $\pi_{U_i}^{0,G'}$ is independent of κ_{ij} . The κ_{ij} shifter affects neither demand, marginal costs, nor retailer payoffs. Dependence of $\pi_{U_i}^{0,G'}$ on $(\lambda_{ij}, \rho_{ij})$ is induced through the graph- G' primitives, so these coefficients have no effect under $G_{-(ij)}$. All displayed affine coefficients are mutually independent. Equations (40)–(42) are the *relationship-local affine parameterization for (ij)* .

Fix any product $l^* \in L_{ij}$. Let the selected affine coordinate be either $\tau_{ij} := \lambda_{l^*,ij}$ or $\tau_{ij} := \rho_{l^*,ij}$, and define the retained coordinate plane

$$\mathcal{D}_{ij} := \text{span} \left\{ \frac{\partial}{\partial \tau_{ij}}, \frac{\partial}{\partial \kappa_{ij}} \right\} \subset T_{\boldsymbol{\theta}^{HL}} \Theta^{HL}.$$

These two parameter directions are active under G and inactive under $G_{-(ij)}$.

Proposition F.1 (Equivalence under logit). For fixed (ij) and $\mathbf{x}_{-(ij)}$, maintain the standing assumptions on X_{ij} , $\Delta \boldsymbol{\pi}_{ij, \mathbf{x}_{-(ij)}}$, and its $C^{K_{ij}}$ regularity. Suppose the homogeneous-logit model satisfies the relationship-local affine parameterization, the stated equilibrium and parameter-independence conditions, and the assumption that any additional price-independent terms in the retailer's payoff cancel between G and $G_{-(ij)}$. Then its gain family satisfies (PR). Hence $B_{ij}^1(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}^{HL}) = B_{ij}^2(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}^{HL}) = B_{ij}^3(\mathbf{x}_{-(ij)}; \boldsymbol{\theta}^{HL})$ for Lebesgue-almost every $\boldsymbol{\theta}^{HL} = (\boldsymbol{\vartheta}^{HL}, \lambda_{ij}, \rho_{ij}, \kappa_{ij}) \in \Theta^{HL}$.

Proof. For each $r \in \mathcal{J}$, define

$$Z_r^{G'} := \sum_{l \in L_r} \exp(v_l^{G'} - \alpha c_l^{G'}) > 0.$$

At an interior price equilibrium, every product of retailer r has markup $[\alpha(1 - \sigma_r^{G'})]^{-1}$. Hence

$$\frac{\sigma_r^{G'}}{s_0^{G'}} = Z_r^{G'} \exp\left(-\frac{1}{1 - \sigma_r^{G'}}\right).$$

Writing $\boldsymbol{\sigma}^{G'} = (\sigma_r^{G'})_{r \in \mathcal{J}}$ and $s_0^{G'} = 1 - \sum_{r \in \mathcal{J}} \sigma_r^{G'}$, define the share system

$$\Phi_r^{G'}(\boldsymbol{\sigma}^{G'}) := \log \sigma_r^{G'} - \log s_0^{G'} + \frac{1}{1 - \sigma_r^{G'}} - \log Z_r^{G'} = 0, \quad r \in \mathcal{J}. \quad (43)$$

Let $\boldsymbol{\Phi}^{G'} := (\Phi_r^{G'})_{r \in \mathcal{J}}$. Its Jacobian with respect to $\boldsymbol{\sigma}^{G'}$ is

$$D_{\boldsymbol{\sigma}} \boldsymbol{\Phi}^{G'} = \text{diag}_{r \in \mathcal{J}} \left(\frac{1}{\sigma_r^{G'}} + \frac{1}{(1 - \sigma_r^{G'})^2} \right) + \frac{1}{s_0^{G'}} \mathbf{1}_J \mathbf{1}_J^\top. \quad (44)$$

This matrix is positive definite: its diagonal term is positive and its rank-one term is positive semidefinite. Thus $(D_{\boldsymbol{\sigma}} \boldsymbol{\Phi}^{G'})^{-1}$ is positive definite and, in particular, $[(D_{\boldsymbol{\sigma}} \boldsymbol{\Phi}^{G'})^{-1}]_{jj} > 0$.

Define

$$\varpi_{l^*|j}^G := \frac{\exp(v_{l^*}^G - \alpha c_{l^*}^G)}{Z_j^G} > 0,$$

the within-retailer exponential weight of product l^* . Changing the selected coordinate τ_{ij} by one changes $v_{l^*}^G - \alpha c_{l^*}^G$ by

$$a_\tau := \begin{cases} 1, & \tau_{ij} = \lambda_{l^*,ij}, \\ -\alpha, & \tau_{ij} = \rho_{l^*,ij}, \end{cases} \quad a_\tau \neq 0,$$

and has no effect under $G_{-(ij)}$. Consequently,

$$\partial_{\tau_{ij}} \log Z_r^G = \mathbf{1}[r = j] a_\tau \varpi_{l^*|j}^G, \quad \partial_{\tau_{ij}} \log Z_r^{G_{-(ij)}} = 0.$$

Differentiating (43) gives

$$\partial_{\tau_{ij}} \sigma_j^G = [(D_{\boldsymbol{\sigma}} \boldsymbol{\Phi}^G)^{-1}]_{jj} a_\tau \varpi_{l^*|j}^G \neq 0, \quad \partial_{\tau_{ij}} \sigma_j^{G_{-(ij)}} = 0.$$

Since $V' > 0$,

$$\partial_{\tau_{ij}} \Delta \pi_{D_j} = V'(\sigma_j^G) [(D_{\boldsymbol{\sigma}} \boldsymbol{\Phi}^G)^{-1}]_{jj} a_\tau \varpi_{l^*|j}^G \neq 0.$$

The supplier shifter κ_{ij} does not affect retailer payoffs, so $\partial_{\kappa_{ij}} \Delta \pi_{D_j} = 0$. Moreover,

$$\partial_{\kappa_{ij}} \Delta \pi_{U_i} = -1,$$

because the relationship cost is present under G and absent under $G_{-(ij)}$. In the ordered coordinate basis $(\partial/\partial \tau_{ij}, \partial/\partial \kappa_{ij})$, the restricted derivative $D_{\mathcal{D}_{ij}} \boldsymbol{\Delta \pi}_{ij, \mathbf{x}_{-(ij)}}$ is represented by

$$\begin{pmatrix} \partial_{\tau_{ij}} \Delta \pi_{D_j} & \partial_{\kappa_{ij}} \Delta \pi_{D_j} \\ \partial_{\tau_{ij}} \Delta \pi_{U_i} & \partial_{\kappa_{ij}} \Delta \pi_{U_i} \end{pmatrix} = \begin{pmatrix} V'(\sigma_j^G) [(D_{\boldsymbol{\sigma}} \boldsymbol{\Phi}^G)^{-1}]_{jj} a_\tau \varpi_{l^*|j}^G & 0 \\ * & -1 \end{pmatrix}.$$

Both diagonal entries are nonzero, so $D_{\mathcal{D}_{ij}} \boldsymbol{\Delta \pi}_{ij, \mathbf{x}_{-(ij)}}$ is onto throughout the regular domain. Apply [Theorem F.1](#) with $\Theta^\circ = \Theta^{HL}$. ■

Remark F.2 (Arbitrarily small affine effects). The rank calculation holds at every admissible value of the affine coefficients. If the zero affine block belongs to the regularity domain, then for every

$\varepsilon > 0$ the domain may be restricted to

$$\Theta_\varepsilon^{HL} := \Theta^{HL} \cap \left\{ \boldsymbol{\theta}^{HL} : \|(\lambda_{ij}, \rho_{ij}, \kappa_{ij})\| < \varepsilon \right\}.$$

Proper parameterization still holds there. The argument therefore requires only arbitrarily small relationship-local perturbations of the primitives.

F.3 Local transfer to mixed logit

The final step uses the logit verification as an anchor. Mixed logit contains homogeneous logit at particular parameter values, which lie on the boundary under standard variance or scale parameterizations. Berry et al. (1999) is a leading example. Let Θ^{ML} be a finite-dimensional $C^{K_{ij}}$ manifold with boundary or corners, and let $\boldsymbol{\theta} \in \Theta^{ML}$ denote the full parameter vector, including all demand, cost, payoff, and heterogeneity parameters. Assume that the relationship-local affine blocks $(\lambda_{ij}, \rho_{ij}, \kappa_{ij})$ remain unrestricted Euclidean coordinates in the relevant boundary charts. Assume also that a $C^{K_{ij}}$ embedding maps Θ^{HL} diffeomorphically onto the relevant boundary face of Θ^{ML} and preserves these affine coordinates. Retain the coordinate plane $\mathcal{D}_{ij}$ defined in the logit construction. The contracting graph, product sets, and ownership structure remain fixed across Θ^{ML} .

For fixed (ij) and $\mathbf{x}_{-(ij)}$, retain the arbitrary-price payoff notation $\pi_f^{G'}(\mathbf{p}, \mathbf{x}^{G'}; \boldsymbol{\theta})$ and the selected equilibrium maps $\mathbf{p}^{*,G'}(\mathbf{x}^{G'}, \boldsymbol{\theta})$ defined above. When using the common domain $X_{ij} \times \Theta^{ML}$, extend every $G_{-(ij)}$ object constantly in the dummy coordinate $\mathbf{x}_{ij}$. If the standing uniqueness requirement in Assumption 1 is relaxed, all payoffs, gains, and bargaining sets in this subsection are understood relative to the fixed selected equilibrium branch. The required smoothness of that branch is imposed explicitly below. The smoothness of mixed-logit demand alone does not guarantee that this composed payoff is smooth: the selected price equilibrium must also vary smoothly. We therefore state both requirements explicitly below.

Corollary F.1 (Local mixed-logit transfer). Let $\boldsymbol{\theta}_0 \in \Theta^{ML}$ correspond to a homogeneous-logit model satisfying the assumptions of Proposition F.1, with the same retained coordinate plane $\mathcal{D}_{ij}$. Suppose there is a common neighborhood $\mathcal{U}$ of $X_{ij} \times \{\boldsymbol{\theta}_0\}$ that is open in $\mathbb{R}^{K_{ij}} \times \Theta^{ML}$ and such that, for each $G' \in \{G, G_{-(ij)}\}$:

1. the selected pricing-equilibrium map $\mathbf{p}^{*,G'}(\mathbf{x}^{G'}, \boldsymbol{\theta})$ is $C^{K_{ij}}$ up to the parameter boundary on $\mathcal{U}$ and is an interior price solution throughout $\mathcal{U}$;
2. there is a neighborhood $\mathcal{W}_{G'}$ of the graph

$$\{(\mathbf{p}^{*,G'}(\mathbf{x}^{G'}, \boldsymbol{\theta}), \mathbf{x}_{ij}, \boldsymbol{\theta}) : (\mathbf{x}_{ij}, \boldsymbol{\theta}) \in \mathcal{U}\}$$

that is open in $\mathbb{R}^L \times \mathbb{R}^{K_{ij}} \times \Theta^{ML}$, on which the mixed-logit demand map and the arbitrary-price payoff maps $\pi_f^{G'}$ are $C^{K_{ij}}$ in $(\mathbf{p}, \mathbf{x}_{ij}, \boldsymbol{\theta})$ up to the parameter boundary; and

3. at $\boldsymbol{\theta}_0$, the derivatives of the induced gain map along $\mathcal{D}_{ij}$ agree with those of the homogeneous-logit model for every $\mathbf{x}_{ij} \in X_{ij}$.

Then there is a relative neighborhood $\mathcal{V} \subset \Theta^{ML}$ of θ_0 on which $D_{\mathcal{D}_{ij}} \Delta \pi_{ij, \mathbf{x}_{-(ij)}}(\mathbf{x}_{ij}, \theta)$ has rank two for every $\mathbf{x}_{ij} \in X_{ij}$ and $\theta \in \mathcal{V}$. Consequently,

$$B_{ij}^1(\mathbf{x}_{-(ij)}; \theta) = B_{ij}^2(\mathbf{x}_{-(ij)}; \theta) = B_{ij}^3(\mathbf{x}_{-(ij)}; \theta)$$

for chart-Lebesgue-almost every mixed-logit parameter $\theta \in \mathcal{V}$.

Proof. The first two assumptions and the composition displayed above imply that the induced payoffs, and hence $\Delta \pi_{ij, \mathbf{x}_{-(ij)}}$, are $C^{K_{ij}}$ up to the parameter boundary on $\mathcal{U}$. The third assumption and [Proposition F.1](#) imply that the derivative along $\mathcal{D}_{ij}$ is onto at θ_0 for every $\mathbf{x}_{ij} \in X_{ij}$.

Define the rank-margin function

$$\text{rm}_{ij}(\mathbf{x}_{ij}, \theta) := \lambda_{\min} \left(D_{\mathcal{D}_{ij}} \Delta \pi_{ij, \mathbf{x}_{-(ij)}}(\mathbf{x}_{ij}, \theta) \left[D_{\mathcal{D}_{ij}} \Delta \pi_{ij, \mathbf{x}_{-(ij)}}(\mathbf{x}_{ij}, \theta) \right]^\top \right),$$

where both restricted derivatives are evaluated at $(\mathbf{x}_{ij}, \theta)$. This function is continuous and positive on the compact set $X_{ij} \times \{\theta_0\}$, so its minimum there is some $r_{ij} > 0$. Since $\mathcal{U}$ is a neighborhood of that compact set, continuity and a finite-cover argument give a relative neighborhood $\mathcal{V}$ of θ_0 such that $X_{ij} \times \mathcal{V} \subset \mathcal{U}$ and $\text{rm}_{ij}(\mathbf{x}_{ij}, \theta) > r_{ij}/2$ throughout $X_{ij} \times \mathcal{V}$. Shrink $\mathcal{V}$ to lie in a parameter chart. If θ_0 is an interior point, choose $\mathcal{V} \subset \text{int } \Theta^{ML}$ and set $\mathcal{V}_{\text{int}} := \mathcal{V}$. Otherwise, define the interior part

$$\mathcal{V}_{\text{int}} := \mathcal{V} \cap \text{int } \Theta^{ML}$$

in the boundary chart. In either case, the chosen chart identifies $\mathcal{V}_{\text{int}}$ with an open subset of $\mathbb{R}^{\dim(\Theta^{ML})}$, and $D_{\mathcal{D}_{ij}} \Delta \pi_{ij, \mathbf{x}_{-(ij)}}$ has rank two throughout $X_{ij} \times \mathcal{V}_{\text{int}}$. Pull back the gain family through this chart. The pulled-back family is $C^{K_{ij}}$ and properly parameterized on that open Euclidean coordinate domain, so [Theorem F.1](#) applies there. Translating the conclusion back through the chart gives the almost-everywhere result on $\mathcal{V}_{\text{int}}$. The boundary of a finite-dimensional manifold with boundary or corners has chart-Lebesgue measure zero, so the same almost-everywhere conclusion holds on all of $\mathcal{V}$. ■

When demand has an integral representation, validity of differentiation under the integral through order K_{ij} is one way to verify the $C^{K_{ij}}$ demand condition in item 2. Cussen (2026) uses the conditions of Morrow and Skerlos (2010a) to show that differentiation under the integral is valid in linear utility models. Cussen (2026) also shows interiority of equilibria and convergence of derivatives to logit derivatives.

On a larger parameter domain, [Theorem F.1](#) gives the same almost-everywhere conclusion whenever the enlarged family is properly parameterized; rank two of the retained block throughout the domain is a simple sufficient condition.

Nesting alone does not establish properness on an arbitrary global parameter domain. Even with analytic maps, pointwise nonvanishing of a minor at logit does not control the projection of its zeros over a continuum of contracts. A global result therefore requires family-level properness or a separate uniform rank argument. In the local result, the homogeneous-logit model is only a boundary anchor for the uniform rank estimate; transversality is applied on the interior, and the parameter boundary is a null set.

G Unique Bargaining Solution with Logit Demand

Let $I = J = 2$, $G = \{(1, 1), (2, 2)\}$. The demand is logit with linear demand and an outside good. Each retailer sells one good, which is the only input it buys from its supplier. The mean utility of good j (identified by the retailer that sells it) is $\delta_j - \alpha p_j$, $\alpha > 0$. Each supplier has marginal cost $c_i > 0$. If bargaining breaks down firms get zero profits. For each (ij) $X_{ij} = [c_i, \bar{x}]$, where $\bar{x} \gg \max\{c_1, c_2\}$ (see Lemma 3). Retailers' only cost is the input price they pay, x_j .

From Konovalov and Sándor (2010) we know that, given $\mathbf{x}$, there exists a unique Bertrand-Nash equilibrium in the second stage game, $\mathbf{p}^*(\mathbf{x})$. Let s_j be the logit choice probability/market share of good j and define $s_j(\mathbf{p}^*(\mathbf{x})) =: s_j^*(\mathbf{x})$.

Proposition 7. *For all $\mathbf{b} \in (0, 1)^2$, for all $(ij) \in G$, for all $\mathbf{x}_{-(ij)} \in X_{-(ij)}$, the Nash product has a unique global maximizer in $(c_i, \bar{x}_i)$.*

Proof. First, note that $(c_i, \bar{x}) \subset B_{ij}(\mathbf{x}_{-(ij)})$ for all (ij) for all $\mathbf{x}_{-(ij)}$, so A2.2 always holds. Fix j and $\mathbf{x}_{-(ij)}$, and suppress subscripts i and j whenever there is no ambiguity. Let $s := s_j^*(\mathbf{x})$. From Appendix F we know that

$$\Delta\pi_{D_j}(\mathbf{x}) = \hat{\pi}_{D_j}(\mathbf{x}) = \frac{s}{\alpha(1-s)}.$$

Let

$$y := \alpha(x - c), \quad t := \frac{s}{1-s} \Rightarrow \Delta\pi_{D_j}(U_i) = \hat{\pi}_{U_i}(\mathbf{x}) = \frac{yt}{\alpha(1+t)} \text{ and } \hat{\pi}_{D_j}(\mathbf{x}) = \frac{t}{\alpha}.$$

Because $\alpha > 0$ we can rewrite the Nash product as

$$N_{ij}(\mathbf{x}, b_{ij}) = \frac{t \cdot y^{1-b_{ij}}}{(1+t)^{1-b_{ij}}}$$

Now we want y as a function of t . Let s_0^* be the share of the outside good in equilibrium given $\mathbf{x}$, and define

$$z := \frac{s_0^*}{1-s_j^*}$$

Then, the shares of all goods in equilibrium are

$$s_j^* = \frac{t}{1+t}, \quad s_0^* = \frac{z}{1+t}, \quad s_k^* = \frac{1-z}{1+t}.$$

For simplicity, think of $j = 1$ and $k = 2$. Note that

$$\log\left(\frac{1-z}{z}\right) = \delta_k - \alpha p_k. \quad (45)$$

The FOC of the logit model is

$$p_k(\mathbf{x}) - x_k = \frac{1}{\alpha(1-s_k)} \Rightarrow \alpha x_k = \alpha p_k + \frac{1}{1-s_k} \quad (46)$$

Thus we get

$$\log\left(\frac{1-z}{z}\right) + \frac{1+t}{t+z} = \delta_k - \alpha x_k. \quad (47)$$

For every $t > 0$, the LHS of (47) is strictly decreasing in z and goes to $+\infty$ as $z \downarrow 0$ and to $-\infty$ as $z \uparrow 1$. It thus defines a smooth function $z = z(t)$ with range $(0, 1)$. Getting its derivative w.r.t. t via implicit differentiation we get

$$z'(t) = -\frac{z(1-z)^2}{D(t, z)} < 0, \quad \text{because} \quad D(t, z) := (t+z)^2 + z(1-z)(1+t) > 0. \quad (48)$$

From the definition of market shares and (46) we have

$$\frac{s_j}{s_0} = \frac{t}{z} = \exp(\delta_j - \alpha p_j) \quad \text{and} \quad \alpha(p_j - x_j) = 1 + t$$

Therefore

$$\phi(t) := t + \ln t - \ln z(t) = \frac{s_j}{1 - s_j} + \delta_j - \alpha p_j \quad (49)$$

Since there is a bijection between i and j , write c_j for the marginal cost of U_i that sells to D_j . Letting $d := \delta_j - \alpha c_j - 1$ we get

$$\begin{aligned} d - \phi(t) &= \delta_j - \alpha c_j - 1 - \frac{s_j}{1 - s_j} - \delta_j + \alpha p_j \\ &= \alpha(p_j - c_j) - \alpha(p_j - x_j) \quad (\text{from the logit FOC}) \\ &= \alpha(x_j - c_j) \end{aligned}$$

But

$$\phi'(t) = 1 + \frac{1}{t} + w(t) > 0 \quad \text{with} \quad \frac{1}{t^2} > w(t) = \frac{(1-z)^2}{D(t, z)} > 0.$$

Thus $y(t)$ is strictly decreasing in t . Therefore, t is strictly decreasing in x_j . Because A2.2 holds, we can take logs inside $B_{ij}(x_k)$ and the FOC of the bargaining problem to get

$$l(t) = \frac{\partial \ln N_{ij}}{\partial x_j} = (1 - b_{ij}) \ln(y(t)) + \ln(t) - (1 - b_{ij}) \ln(1 + t)$$

Because $y'(t) = -\phi'(t)$ we get

$$l'(t) = -\frac{(1 - b_{ij})\phi'(t)}{y(t)} + \frac{1 + b_{ij}t}{t(1 + t)} \quad (50)$$

Define

$$A_{b_{ij}}(t) := \frac{(1 - b_{ij})t(1 + t)}{1 + b_{ij}t}, \quad K_{b_{ij}}(t) := \phi(t) + A_{b_{ij}}(t)\phi'(t). \quad (51)$$

Equations (49) and (50) imply

$$\text{sgn } \ell'(t) = \text{sgn}[d - K_{b_{ij}}(t)]. \quad (52)$$

Now we want to prove that $K_{b_{ij}}$ is strictly increasing. Begin by differentiating w along the function $z = z(t)$. This gives

$$\frac{w'}{w} = -\frac{2t + z + z^2}{D} + \frac{z(1-z)^2(3t + 1 - 2tz)}{D^2}.$$

The second term in the RHS is positive, and

$$2D - t(2t + z + z^2) = z[2 + t(5 - 3z)] > 0.$$

Thus

$$w'(t) > -\frac{2w(t)}{t}. \quad (53)$$

Define

$$\phi'_0(t) := 1 + \frac{1}{t}, \quad \phi''_0(t) := -\frac{1}{t^2},$$

so that $\phi' = \phi'_0 + w$ and $\phi'' = \phi''_0 + w'$. Differentiating $K_{b_{ij}}$ gives

$$K'_{b_{ij}} = K'_{b_{ij},0} + w(1 + A'_{b_{ij}}) + A_{b_{ij}}w', \quad (54)$$

where

$$\begin{aligned} K'_{b_{ij},0} &:= (1 + A'_{b_{ij}})\phi'_0 + A_{b_{ij}}\phi''_0 \\ &= \frac{(1+t)[b_{ij}t^2 + ((1-b_{ij})^2 - (1-b_{ij}) + 2)t + 1]}{t[1+b_{ij}t]^2} > 0. \end{aligned} \quad (55)$$

Let

$$B_{b_{ij}}(t) := 1 + A'_{b_{ij}}(t) - \frac{2A_{b_{ij}}(t)}{t}. \quad (56)$$

By (53),

$$w(1 + A'_{b_{ij}}) + A_{b_{ij}}w' > wB_{b_{ij}}. \quad (57)$$

If $B_{b_{ij}} \geq 0$, equations (54), (55), and (57) imply $K'_\beta > 0$.

Suppose instead that $B_{b_{ij}} < 0$. Since $w < 1/t^2$, multiplying this inequality by the negative number $B_{b_{ij}}$ reverses its direction, so

$$wB_{b_{ij}} > \frac{B_{b_{ij}}}{t^2}.$$

It follows that

$$K'_{b_{ij}} > K'_{b_{ij},0} + \frac{B_{b_{ij}}}{t^2} = \frac{(1+t)P_{b_{ij}}(t)}{t^2[1+b_{ij}t]^2}, \quad (58)$$

where

$$P_{b_{ij}}(t) = b_{ij}t^3 + ((1-b_{ij})^2 - (1-b_{ij}) + 2)t^2 + (2(1-b_{ij})^2 - 3(1-b_{ij}) + 2)t + b_{ij}. \quad (59)$$

Every coefficient in (59) is strictly positive because $b_{ij} \in (0, 1)$. Indeed,

$$(1-b_{ij})^2 - (1-b_{ij}) + 2 = \left((1-b_{ij}) - \frac{1}{2}\right)^2 + \frac{7}{4} > 0$$

and

$$2(1-b_{ij})^2 - 3(1-b_{ij}) + 2 = 2\left((1-b_{ij}) - \frac{3}{4}\right)^2 + \frac{7}{8} > 0.$$

Therefore $P_{b_{ij}}(t) > 0$ for every $t > 0$. Together with the preceding case this proves that

$$K'_\beta(t) > 0 \quad \text{for every } t > 0. \quad (60)$$

By (52) and (60), $l'(t)$ can equal zero at most once. If it equals zero, it changes sign from positive to negative. Thus the Nash product is strictly single-peaked as a function of t . Since t is strictly decreasing in x_j , the same unique maximizer obtains when the objective is viewed as a function of x_j .

At $x_j = c_j$, U_i 's gains from trade and the Nash product are zero. Every jointly profitable contract has a strictly positive Nash product. Therefore, if the unique peak lies in $(c_j, \bar{x})$, it is the unique global maximizer. Now we note that this logit model is a special case of linear utility mixed logit models, to which the mixed logit models converge uniformly as the variances of their random coefficients (in this case, without loss, of α) go to zero (Cussen, 2026). Therefore, proofs for mixed logit models apply here. Corollary 2.18 of Morrow and Skerlos (2010a) shows that eventually the derivatives of π_{D_j} w.r.t. p_j will become negative. From Cussen (2026) we know that that result applies in this model. Since p_j is increasing in x_c , if x_j is sufficiently high we get $l'(t) < 0$ at some $(c_i, \bar{x})$. The existence of $\bar{x}$ high enough follows from Corollary 2.18 of Morrow and Skerlos (2010a). This completes the proof. ■

Comment. If $\mathbf{b} \rightarrow \mathbf{1} \in \mathbb{R}^2$, then Proposition 7 shows that it can be proven analytically that all our assumptions can hold in a single model with discrete choice demand. Table 5 summarizes this result.

Table 5: Verifying Assumptions in a Simple 2×2 Logit

Assumption	
Assumption 1	Konovalov and Sándor (2010)
Assumption 2.1	Assumptions 1 and 5, and Proposition 4
Assumption 2.2	$B_{ij}(\mathbf{x}_{-(ij)}) \supset (c_i, \bar{x}) \forall \mathbf{x}_{-(ij)} \in X_{-(ij)} \forall (ij) \in G$
Assumption 2.3	Proposition 7
Assumption 2.4	Lemma 3
Assumption 3.1	Assumptions 2 and 7 + Proposition 5, or Proposition 6
Assumption 3.2	Lemma 5
Assumption 4	Lemma 7 if $\mathbf{b} \rightarrow \mathbf{1} \in \mathbb{R}^2$
Assumption 5.1	Appendix E
Assumption 5.2	Konovalov and Sándor (2010)
Assumption 6	Appendix E
Assumption 7	Proposition 6

Notes: All our assumptions can be verified directly in a model with 2 single product suppliers, two single product retailers, two bilateral monopoly relations, linear input prices, logit demand, and zero profits if bargaining breaks down.

References

- Aksoy-Pierson, M., G. Allon, and A. Federgruen (2013). Price competition under mixed multinomial logit demand functions. *Management Science* 59(8), 1817–1835.
- Albæk, S., P. Møllgaard, and P. B. Overgaard (1997). Government-assisted oligopoly coordination? A concrete case. *The Journal of Industrial Economics* 45(4), 429–443.
- Alviarez, V. I., M. Fioretti, K. Kikkawa, and M. Morlacco (2023). Two-sided market power in firm-to-firm trade. *NBER Working Paper* 31253.
- Asil, A., P. Ramos, A. Starc, and T. G. Wollmann (2026). Painful bargaining: Evidence from anesthesia rollups. *NBER Working Paper* 33217.
- Asker, J. (2016). Diagnosing foreclosure due to exclusive dealing. *The Journal of Industrial Economics* 64(3), 375–410.
- Asker, J. and H. Bar-Isaac (2014). Raising Retailers’ Profits: On Vertical Practices and the Exclusion of Rivals. *American Economic Review* 104(2), 672–686.
- Bagwell, K., R. W. Staiger, and A. Yurukoglu (2020). “Nash-in-Nash” tariff bargaining. *Journal of International Economics* 122, 103263.
- Bagwell, K., R. W. Staiger, and A. Yurukoglu (2021). Quantitative analysis of multiparty tariff negotiations. *Econometrica* 89(4), 1595–1631.
- Barwick, P. J., A. T. Swanson, and T. Xia (2025). A Double Dose of Reform: Insurance and Centralized Negotiation in Drug Markets. *NBER Working Paper* 33832.
- Beckert, W., H. W. Smith, and Y. Takahashi (2020). Competition in a spatially-differentiated product market with negotiated prices. *CEPR Discussion Paper No. DP15379*.
- Bernheim, B. D., B. Peleg, and M. D. Whinston (1987). Coalition-proof Nash equilibria I. Concepts. *Journal of Economic Theory* 42(1), 1–12.
- Berry, S., J. Levinsohn, and A. Pakes (1999). Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy. *American Economic Review* 89(3), 400–431.
- Bhattacharya, V. and G. Illanes (2022). The design of defined contribution plans. *NBER Working Paper* 29981.
- Borkovsky, R. N., U. Doraszelski, and Y. Kryukov (2010). A User’s Guide to Solving Dynamic Stochastic Games Using the Homotopy Method. *Operations Research* 58(4), 1116–1132.
- Brekke, K. R., T. H. Holmås, and O. R. Straume (2015). Price regulation and parallel imports of pharmaceuticals. *Journal of Public Economics* 129(1), 92–105.
- Cachon, G. P. (2003). Supply chain coordination with contracts. In *Handbook Operations Research and Management Science*, Volume 11, pp. 229–339. Elsevier.
- Calzolari, G., V. Denicolo, and P. Zanchettin (2020). The Demand-boost Theory of Exclusive Dealing. *The RAND Journal of Economics* 51(3), 713–738.
- Caplin, A. and B. Nalebuff (1991). Aggregation and Imperfect Competition: On the Existence of Equilibrium. *Econometrica* 59(1), 25–59.

- Caprice, S. and V. von Schlippenbach (2013). One-stop shopping as a cause of slotting fees: A rent-shifting mechanism. *Journal of Economics & Management Strategy* 22(3), 468–487.
- Chambolle, C. and H. Molina (2023). A Buyer Power Theory of Exclusive Dealing and Exclusionary Bundling. *American Economic Journal: Microeconomics* 15(3), 166–200.
- Chipty, T. and C. M. Snyder (1999). The role of firm size in bilateral bargaining: A study of the cable television industry. *Review of Economics and Statistics* 81(2), 326–340.
- Collard-Wexler, A., G. Gowrisankaran, and R. S. Lee (2019). “Nash-in-Nash” Bargaining: A Micro-foundation for Applied Work. *Journal of Political Economy* 127(1), 163–195.
- Conlon, C. and J. Gortmaker (2020). Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics* 51(4), 1108–1161.
- Crawford, G. S., R. S. Lee, M. D. Whinston, and A. Yurukoglu (2018). The Welfare Effects of Vertical Integration in Multichannel Television Markets. *Econometrica* 86(3), 891–954.
- Crawford, G. S. and A. Yurukoglu (2012). The Welfare Effects of Bundling in Multichannel Television Markets. *American Economic Review* 102(2), 643–685.
- Cuesta, J. I., C. E. Noton, and B. Vatter (2025). Vertical integration and plan design in healthcare markets. *NBER Working Paper* 32833.
- Cussen, D. (2026). Unique Equilibrium in Models with Mixed Logit Demand Under Variance Constraints. *Working Paper*.
- Cussen, D. and J.-P. Montero (2022). Using list prices to collude or to compete? *Working Paper, IE-PUC* 567.
- Cussen, D. and J.-P. Montero (2024). Using list prices to collude or to compete? *Economic Journal* 134(664), 3232–3261.
- Demirer, M. and M. Rubens (2025). Welfare Effects of Buyer and Seller Power. *Working Paper*.
- Dix, R. and T. Lensman (2025). Combining complements: Theory and evidence from cancer treatment innovation. *Working Paper*.
- Dobson, P. W. and M. Waterson (1997). Countervailing power and consumer prices. *Economic Journal* 107, 418–430.
- Dubois, P. and M. Sæthre (2020). On the Effect of Parallel Trade on Manufacturers’ and Retailers’ Profits in the Pharmaceutical Sector. *Econometrica* 88(6), 2503–2545.
- Elberg, A. and C. Noton (2024). What do trade allowances look like? Evidence from actual payments to a big-box retailer. *Management Science, Forthcoming*.
- Escrhuella-Villar, M., W. Ferrarese, A. Iozzi, and J. Tolvanen (2024). Bargaining power and Nash-in-Nash bargaining: When more is less. *Working Paper*.
- Galichon, A. (2016). *Optimal Transport Methods in Economics*. Princeton: Princeton University Press.
- Galichon, A. and M. Henry (2011). Set identification in models with multiple equilibria. *Review of Economic Studies* 78(4), 1264–1298.
- Galichon, A., L. Samuelson, and L. Vernet (2025). Unified gross substitutes and inverse isotonicity for equilibrium problems. *Theoretical Economics* 20(3), 941–971.

- Gallego, G., W. T. Huh, W. Kang, and R. Phillips (2006). Price competition with the attraction demand model: Existence of unique equilibrium and its stability. *Manufacturing & Service Operations Management* 8(4), 359–375.
- Gallego, G. and R. Wang (2014). Multiproduct Price Optimization and Competition Under the Nested Logit Model with Product-Differentiated Price Sensitivities. *Operations Research* 62(2), 450–461.
- Gandhi, A. and A. Nevo (2021). Empirical models of demand and supply in differentiated products industries. In *Handbook of Industrial Organization*, Volume 4, pp. 63–139. Elsevier.
- Garrido, F. (2024). An aggregative approach to pricing equilibrium among multi-product firms with nested demand. *The RAND Journal of Economics* 55(3), 359–374.
- Gemulla, R., E. Nijkamp, P. J. Haas, and Y. Sismanis (2011). Large-scale matrix factorization with distributed stochastic gradient descent. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 69–77.
- Gilbert, R. J. (2015). E-books: A Tale of Digital Disruption. *Journal of Economic Perspectives* 29(3), 165–184.
- González, G. and S. Poblete (2024). Union bargaining schemes in the presence of market power.
- Gowrisankaran, G., A. Nevo, and R. Town (2015). Mergers When Prices are Negotiated: Evidence from the Hospital Industry. *American Economic Review* 105(1), 172–203.
- Grennan, M. (2013). Price Discrimination and Bargaining: Empirical Evidence from Medical Devices. *American Economic Review* 103(1), 145–177.
- Grennan, M. (2014). Bargaining Ability and Competitive Advantage: Empirical Evidence from Medical Devices. *Management Science* 60(12), 3011–3025.
- Grennan, M. and A. Swanson (2020). Transparency and negotiated prices: The value of information in hospital-supplier bargaining. *Journal of Political Economy* 128(4), 1234–1268.
- Gross, A. and D. Savelle (2025). Alternative Contract Frameworks in Nash-in-Nash Bargaining Models. *Working Paper*.
- Guillemin, V. and A. Pollack (1974). *Differential Topology*. Englewood Cliffs, NJ: Prentice Hall.
- Hastings, J. S. and R. J. Gilbert (2005). Market power, vertical integration and the wholesale price of gasoline. *The Journal of Industrial Economics* 53(4), 469–492.
- Herrero, M. J. (1989). The Nash Program: Non-convex Bargaining Problems. *Journal of Economic Theory* 49(2), 266–277.
- Hirsch, M. W. (2012). *Differential Topology*, Volume 33. Springer Science & Business Media.
- Ho, K. and R. S. Lee (2017). Insurer competition in health care markets. *Econometrica* 85(2), 379–417.
- Ho, K. and R. S. Lee (2019). Equilibrium Provider Networks: Bargaining and Exclusion in Health Care Markets. *American Economic Review* 109(2), 473–522.
- Horn, H. and A. Wolinsky (1988). Bilateral Monopolies and Incentives for Merger. *The RAND Journal of Economics* 19(3), 408–419.
- Horn, R. A. and C. R. Johnson (2013). *Matrix Analysis*. Cambridge: Cambridge University Press.
- Hristakeva, S. (2022). Vertical contracts with endogenous product selection: An empirical analysis of

- vendor allowance contracts. *Journal of Political Economy* 130(12), 3202–3252.
- Iaria, A. and A. Wang (2019). Identification and Estimation of Demand for Bundles. *Working Paper*.
- Ide, E. and J.-P. Montero (2024). Monopolization with must-haves. *American Economic Journal: Microeconomics* 16(3), 284–320.
- Iyer, G. and J. M. Villas-Boas (2003). A bargaining theory of distribution channels. *Journal of Marketing Research* 40(1), 80–100.
- Jovanovic, B. (1989). Observable Implications of Models with Multiple Equilibria. *Econometrica* 57(6), 1431–1437.
- Kadiyali, V., P. Chintagunta, and N. Vilcassim (2000). Manufacturer-retailer channel interactions and implications for channel power: An empirical investigation of pricing in a local market. *Marketing Science* 19(2), 127–148.
- Kononov, A. and Z. Sándor (2010). On price equilibrium with multi-product firms. *Economic Theory* 44(2), 271–292.
- Koren, Y., R. Bell, and C. Volinsky (2009). Matrix factorization techniques for recommender systems. *IEEE Computer* 42(8), 30–37. Seminal paper on large-scale collaborative filtering using gradient-descent matrix factorization (Netflix Prize benchmark).
- Kourandi, F. and I. N. Pinopoulos (2022). Vertical integration and bargaining: Linear vs. two-part tariffs. *Working Paper*.
- Lee, J. D., M. Simchowitz, M. I. Jordan, and B. Recht (2016, 23–26 Jun). Gradient Descent Only Converges to Minimizers. In V. Feldman, A. Rakhlin, and O. Shamir (Eds.), *29th Annual Conference on Learning Theory*, Volume 49 of *Proceedings of Machine Learning Research*, Columbia University, New York, New York, USA, pp. 1246–1257. PMLR.
- Lee, R. S. and K. Fong (2013). Markov Perfect Network Formation: An Applied Framework for Bilateral Oligopoly and Bargaining in Buyer-Seller Networks. *Discussion Paper, Harvard Univ.*
- Lee, R. S., M. D. Whinston, and A. Yurukoglu (2021). Structural Empirical Analysis of Contracting in Vertical Markets. In *Handbook of Industrial Organization*, Volume 4, pp. 673–742. Elsevier.
- Liu, G. (2006). On Nash equilibrium in prices in an oligopolistic market with demand characterized by a nested multinomial logit model and multiproduct firm as nest. *Statistics Norway*.
- Luco, F. and G. Marshall (2020). The competitive impact of vertical integration by multiproduct firms. *American Economic Review* 110(7), 2041–2064.
- Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. Oxford: Oxford University Press.
- Meza, S. and K. Sudhir (2010). Do private labels increase retailer bargaining power? *Quantitative Marketing and Economics* 8(3), 333–363.
- Miller, N. H. and M. C. Weinberg (2017a). Supplement to ‘Understanding the price effects of the MillerCoors joint venture’. *Econometrica Supplemental Material*, 85, <http://dx.doi.org/10.3982/ECTA13333>. [1766,1780].
- Miller, N. H. and M. C. Weinberg (2017b). Understanding the price effects of the MillerCoors joint

- venture. *Econometrica* 85(6), 1763–1791.
- Milnor, J. W. (1965). *Topology from the Differentiable viewpoint*. University Press of Virginia.
- Mnih, A. and R. R. Salakhutdinov (2007). Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems (NeurIPS)*, Volume 20, pp. 1257–1264.
- Morrow, W. R. (2015). Finite purchasing power and computations of Bertrand-Nash equilibrium prices. *Computational Optimization and Applications* 62(2), 477–515.
- Morrow, W. R. and S. J. Skerlos (2010a). Fixed-point Approaches to Computing Bertrand-Nash Equilibrium Prices under Mixed Logit Demand: A Technical Framework for Analysis and Efficient Computational Methods. *arXiv:1012.5836v1*.
- Morrow, W. R. and S. J. Skerlos (2010b). On the Existence of Bertrand-Nash Equilibrium Prices Under Logit Demand. *arXiv preprint arXiv:1012.5832*.
- Morrow, W. R. and S. J. Skerlos (2011). Fixed-point approaches to computing Bertrand-Nash equilibrium prices under mixed-logit demand. *Operations Research* 59(2), 328–345.
- Moulin, H. (1984). Dominance Solvability and Cournot Stability. *Mathematical Social Sciences* 7(1), 83–102.
- Nehzati, M. and D. Cussen (2026). Inference From Random Restarts. *Working Paper*.
- Nguyen, T. and R. Vohra (2018). Near-feasible stable matchings with couples. *American Economic Review* 108(11), 3154–3169.
- Nocke, V. and N. Schutz (2018). Multiproduct-firm oligopoly: An aggregative games approach. *Econometrica* 86(2), 523–557.
- Nocke, V. and N. Schutz (2025). Oligopoly, complementarities, and transformed potentials.
- Noton, C. and A. Elberg (2018). Are Supermarkets Squeezing Small Suppliers? Evidence from Negotiated Wholesale Prices. *Economic Journal* 128(May), 1304–1330.
- Panhans, M. T. (2024). Vertical integration in a sequential model of a supply chain with bargaining. *FTC Working Paper NO. 349*.
- Rao, A. R. and H. Mahi (2003). The price of launching a new product: Empirical evidence on factors affecting the relative magnitude of slotting allowances. *Marketing Science* 22(2), 246–268.
- Rey, P. and T. Vergé (2020). Multilateral vertical contracts. *Working Paper*.
- Robles-Garcia, C. (2020). Competition and incentives in mortgage markets: The role of brokers.
- Shaffer, G. (1991). Slotting allowances and resale price maintenance: A comparison of facilitating practices. *The RAND Journal of Economics* 22(1), 120–135.
- Sheu, G. and C. Taragin (2021). Simulating mergers in a vertical supply chain with bargaining. *The RAND Journal of Economics* 52(3), 596–632.
- Simsek, A., A. Ozdaglar, and D. Acemoglu (2007). Generalized Poincaré-Hopf theorem for compact nonsmooth regions. *Mathematics of Operations Research* 32(1), 193–214.
- Stokey, N. L. and R. E. Lucas Jr (1989). *Recursive Methods in Economic Dynamics*.
- Tirole, J. (1988). *The Theory of Industrial Organization*. Cambridge: MIT Press.
- Zheng, H. (2024). Death by Market Power: The Production-Safety Tradeoff in the Coal Mining Industry.