On Stochastic Games, Reinforcement Learning, and Platform Competition and Collusion

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Cristian Camilo Chica Castano

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Gilad Lerman, Adviser

May 2025

ACKNOWLEDGEMENTS

I would like to express my gratitude to every person who has supported my academic and personal journey.

I am indebted to my advisor, Dr. Gilad Lerman, for his continued guidance, patience, and kindness. The countless hours he dedicated to improving our papers are a testament to his commitment to each of his students. My most sincere appreciation goes to the members of my dissertation committee, Dr. Jeff Calder, Dr. William Leeb, and Dr. Yulong Lu, for their time and insightful feedback. I am also thankful to the School of Mathematics, including both staff members and faculty. Special thanks to my co-author and fellow Dr. Yinglong Guo for helping me find an exit whenever I hit a mathematical wall.

To my closest fellows and friends, Devraj Duggal, Sarbartha Bhattacharya, David Luo, and Adriaan de Clercq, thank you for making those first Ph.D. years a marvelous memory.

To my wife, Luisa Fernanda Garcia Chaverra, thank you for your patience throughout all these years and for teaching me so many things I could not learn in school.

To my son, Isaac, thank you for giving me the time to finalize my dissertation.

To my family—Libaniel, Nelsy, Nelly and Francedine—thank you for your continued support.

Finally, to many others: Carlos Augusto Velez Lopez, Jose Manuel Jimenez Urrea, and Luis Fernando Quiroz Jimenez.

The work reported in this dissertation was partially supported by NSF DMS awards 2124913 and 2427955.

ABSTRACT

In this dissertation we formulate a version of Q-learning with bounded experimentation in a setting of stochastic games with bounded memory and show sufficient conditions under which firms learn that charging supracompetitive prices is optimal in the long run. We also show sufficient conditions for these supracompetitive prices to be supported by three types of different strategies known as naive collusion, grim trigger and increasing strategies. Then, we study what is competition and collusion in a static game model of two-sided markets with an outside option. Comparing collusion to competition, we find that in cases of small cross-side externalities, collusion results in decreased normalized net deterministic utilities, reduced market participation and increased price, on both sides of the market. We quantify the effects of different model parameters in the equilibrium quantities and provide a wide range of economic interpretations. Finally, we examine how AI agents using Q-learning engage in tacit collusion in two-sided markets. We show that collusion by these AI-driven agents is feasible under different choice of the model parameters.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this dissertation, we address the question of whether Q-learning-driven firms can learn that charging supracompetitive prices is optimal in the long run. If so, what are some sufficient conditions for this to happen, and what kinds of economic strategies support this behavior? This question was originally posed by Waltman and Kaymak (2008) in the context of Cournot competition, where firms compete by choosing how much to produce at each stage of the game. A limitation of that paper is that it does not explore the dynamics of learning, nor does it explain why Q-learning agents might conclude that collusion is optimal. Later, Calvano et al. (2020b) showed numerically that firms can learn to charge supracompetitive prices (i.e., prices above the competitive level). They also found that agents learn strategies consistent with collusive behavior—for example, they return to the previous pricing strategy after unilateral deviations, and they develop punishment and reward mechanisms. However, their results are based solely on simulations and are specific to Bertrand competition. This dissertation makes three main contributions to the fields of reinforcement learning and game theory: (i) it presents the first theoretical result explaining how collusion can be sustained by Q-learning-driven firms in stochastic games where both a one-stage Nash equilibrium price and a collusive-enabling price exist; (ii) it introduces an outside option feature into a model of platform competition in two-sided markets, enabling the study of collusion in this economic setting; and (iii) it examines how AI agents using

Q-learning engage in tacit collusion in two-sided markets.

The rest of the dissertation is organized as follows. Chapter 2 develops a model for stochastic games with bounded memory and shows the existence of one-memory subgame perfect Nash equilibria (SPEs); it introduces a version of Q-learning with bounded experimentation and shows sufficient conditions for the convergence to supracompetitive prices; it provides sufficient conditions for these convergence to be suported by Naive Collusion, Grim Trigger and Increasing Strategies.

Chapter 2

Stochastic Games, Reinforcement Learning and Collusion with Bounded Memory

2.1 Introduction

Collusion by algorithmic driven firms has been a major topic of discussion in recent years. Since the publication of Calvano et al. (2020b), many papers have addressed, in different economic settings and using different algorithms, whether algorithmic driven firms may learn collusion as an optimal long-run behavior. However, most of the existing studies are limited to numerical experiments and theoretical key questions remain open: (i) Under what conditions do firms learn that supracompetitive prices are optimal in the long run? (ii) Are these supracompetitive prices supported by strategies exhibiting punishment and reward behavior? (iii) Does the learned behavior align with Nash equilibrium behavior? In this paper, we answer these questions. To do so, we develop a framework of stochastic games with bounded memory and study their one-memory subgame perfect Nash equilibria (SPEs). We then formulate a version of Q-learning with bounded experimentation and study the rise of supracompetitive prices and collusion in our stochastic setting.

In our stochastic game model, n firms compete on an infinite time horizon by choosing one-memory policies, i.e., strategies that only depend on immediate past actions (or prices) and the current state. The action and state sets are finite, while states evolve according to a probability law also depending on previous prices and state. Profit functions depend on current vector of prices and state. Moreover, two value functions characterize our stochastic game, the conditional value functions at time $t = 0$ and $t = 1$, measuring the expected discounted payoff of the game from starting from time $t = 0$ or $t = 1$.

By establishing a connection between the conditional value function from time $t = 1$ and a vector valued function from Fink (1964), we leverage his theory of stationary points to establish the existence of one-memory SPEs in our stochastic game setting, and formulate an algorithm for checking whether a given profile is a one-memory SPE. We then apply this algorithm to stochastic games where there is a one-stage Nash equilibrium price, and a collusive-enabling price and show sufficient conditions for grim trigger strategies to be a one-memory SPE.

Next, we formulate a version of the Q-learning algorithm with no experimentation, adapting it to our stochastic setting, and establish a connection between the fixed points of the algorithm and the value functions of the stochastic game. Specifically, we show that these fixed points coincide with the conditional value function of the stochastic game at time $t = 1$ for a specific type of strategies, named induced strategies. We then formulate a version of Q-learning with bounded experimentation and show sufficient conditions for the convergence of this algorithm in stochastic games where there is a one-stage Nash equilibrium price, and a collusive-enabling price. We note that dynamic Bertrand competition satisfies the latter hypotheses. We then show sufficient conditions under which Q-learning firms learn that choosing supracompetitive prices is optimal. We also show sufficient conditions under which these supracompetitive prices are supported by either naive collusion, grim trigger strategies or increasing strategies, and whether they align with Nash equilibrium behavior.

In Q-learning with bounded experimentation, firms play the Q-learning algorithm following a soft-max value over the Q-function, then, at some finite time, they switch to the argmax and play Q-learning with no experimentation. Our sufficient conditions for the convergence of Q-learning with bounded experimentation specify a set of inequalities between the profit function at the collusive-enabling price and the entries of the Q-function at the time in which experimentation stops. Moreover, it specifies a set of inequalities the Q-function must satisfies in order to ensure convergence to supracompetitive prices.

Related Literature. This work is inspired by earlier studies on algorithmic price collusion showing some of the unintended consequences of using multi-agent reinforcement learning (MARL) algorithms in economic settings. For instance, Waltman and Kaymak (2008), Calvano et al. (2020b), Klein (2021) and Chica et al. (2024a) show via numerical experiments that Q-learning agents can learn to collude in multiple economic settings. This collusion harms consumers, thereby raising the alarms of governments and anti-trust institutions (see, e.g., OECD (2017) and Assad et al. (2024)). Our work shows sufficient conditions for Q-learning driven firms to consistently choose the collusive-enabling price in every stage of the repeated game. Thus showing that firms learn to choose supracompetitive prices in the long run.

Theoretical research has shown that simple pricing algorithms can raise prices (Brown and MacKay (2023)). Our work shows sufficient conditions for this raise to happen in the case of Q-learning algorithms. The latter algorithm is well-known and common in reinforcement learning theory, first developed by Watkins and Dayan (1992), is commonly used to a find a maximal policy for a given value function. The convergence of such algorithm is tricky and for a single agent, Jaakkola et al. (1993) provides sufficient conditions for its convergence. Our theory shows sufficient conditions for the convergence of Q-learning by n firms in stochastic games where there is a one-stage Nash equilibrium price, and a collusive-enabling price.

Our work is related to the theory of general-sum stochastic games.[1] We extend Fink

[1]Previous work has focused on finding equilibria in zero-sum stochastic games, where the overall payoff

(1964), where agents use stationary strategies, to the case of one-memory policies. The latter type of policy is more general than stationary strategies and allows us to capture potential punishment and reward behavior by the competing firms. For example, grim trigger type strategies are one-memory policies. Our theory of stochastic games with bounded memory is also related to Barlo et al. (2009), where infinite repeated games are studied. These authors show that if the set of actions is sufficiently rich and agents are sufficiently patient, then any subgame perfect equilibrium can be supported with a one-memory strategy. One of the limitation of this work is that the game is static at every period in time. By contrast, we study stochastic games where a state is updated at every time t following a given probabilistic rule.

To our knowledge, this work contains the first theoretical result explaining how collusion can be sustained by Q-learning driven firms in the context of repeated Bertrand Competition.[2]

2.2 A Model for Stochastic Games with Bounded Memory

In this section, we introduce our stochastic game model. To make the problem more concrete, we assume that n firms compete in prices over an infinite time horizon. Each firm is indexed by $i \in [n] := \{1, \dots, n\}$. Nevertheless, in general one can consider any n agents with an ordered finite set of actions, which are prices for us. We start describing the basic components of this stochastic game and then explain in Section 2.2.1 the condition value function of firm i, that is, the V^i-function.

Actions: We assume a set of actions $\mathcal{A} := \{a^0, \dots, a^m\}$. We recall that in our con-

is balanced to zero, so a loss of one player is a win for other players. In particular, Fink (1964) showed that stationary equilibria always exist in zero-sum stochastic games with finitely many players, states and actions. By contrast, in a general-sum stochastic game, the sum of player's payoffs in each stage of the game may not be zero. The latter game is thus helpful to describe situations in which agents can cooperate, collaborate and compete. The question about existence of equilibria in a general-sum stochastic game is more difficult to analyze, as the value of the game is not uniquely defined (Ozdaglar et al., 2021).

[2] The work by Possnig (2023) shows similar results about collusion for the case of Cournot game.

text taking actions means charging prices. The set of actions for n agents is $\mathcal{A}^n$ and we commonly denote by $\boldsymbol{p} = (p^1, \ldots, p^n)$, a vector of prices in $\mathcal{A}^n$.

States and their dynamics: We assume a state space of r states: $\mathcal{S} := \{s^1, \ldots, s^r\}$. Every state may represent a market demand or cost level, which will directly affect the payoff functions defined below. States change with time and consequently affect the profits agents receive. At time $t+1$, given state $s_t = s \in \mathcal{S}$ and vector of prices $\boldsymbol{p} = (p^1, \ldots, p^n) \in \mathcal{A}^n$, the state at $t+1$, $s_{t+1} \in \mathcal{S}$, follows the probabilistic law

$$s_{t+1} \sim \mathbb{P}(\cdot|\boldsymbol{p}, s). \tag{2.1}$$

Therefore, the state at $t+1$ only depends on the state and price vector at time t.

Profit functions: The profit function for each firm i is a function,

$$\pi^i : \mathcal{A}^n \times \mathcal{S} \to \mathbb{R}. \tag{2.2}$$

We note that it is a function of the current vector of prices, $\boldsymbol{p} = (p^1, \ldots, p^n) \in \mathcal{A}^n$, and state, $s \in \mathcal{S}$, but independent of the time t. Moreover, we assume that $\pi^i \geq 0$. Computer scientists refer to π^i as the reward function, but our terminology corresponds to our economic interpretation.

Policies: A policy, or strategy, for firm i is a sequence of probability distributions $\boldsymbol{\sigma}^i = (\sigma^i_t)_{t=0}^{\infty}$ over the action space $\mathcal{A}$.[3] At time $t = 0$ and given a state $s_0 \in \mathcal{S}$, firm i chooses $p \in \mathcal{A}$ with probability $\sigma^i_0(p|s_0)$, where $\sum_{p\in\mathcal{A}} \sigma^i_0(p|s_0) = 1$. Let p^i_{t-1} denote the price chosen by firm i in period $t-1$ and let $\boldsymbol{p}_{t-1} = (p^1_{t-1}, \ldots, p^n_{t-1}) \in \mathcal{A}^n$ denote the vector of all these prices. We assume that at time $t \geq 1$, $\boldsymbol{p}_{t-1}$ is publicly available. At time $t \geq 1$ and given $s_t \in \mathcal{S}$ and $\boldsymbol{p}_{t-1} \in \mathcal{A}^n$, firm i chooses $p \in \mathcal{A}$ with probability $\sigma^i_t(p|\boldsymbol{p}_{t-1}, s_t)$, where $\sum_{p\in\mathcal{A}} \sigma^i_t(p|\boldsymbol{p}_{t-1}, s_t) = 1$. We assume that $\sigma^1_t(p^1_t|\boldsymbol{p}_{t-1}, s_t), \ldots, \sigma^n_t(p^n_t|\boldsymbol{p}_{t-1}, s_t)$ are

[3] Computer scientists use the terminology policy, whereas economists use the terminology strategy.

independent random variables. Consequently, we define

$$\sigma_t(\boldsymbol{p}_t|\boldsymbol{p}_{t-1},s_t)=\prod_{i=1}^{n}\sigma_t^i(p_t^i|\boldsymbol{p}_{t-1},s_t)\ \text{ and }\ \sigma_t^{-i}(\boldsymbol{p}_t|\boldsymbol{p}_{t-1},s_t)=\prod_{j\neq i}\sigma_t^j(p_t^j|\boldsymbol{p}_{t-1},s_t).$$

We similarly define $\sigma_0(\boldsymbol{p}_0|s_0)$ and $\sigma_0^{-i}(\boldsymbol{p}_0|s_0)$.

We further assume the following key condition, which is common in repeated games with bounded memory (see, e.g., Barlo et al. (2009) and Barlo et al. (2016)):

Assumption 1 (One-memory policies)**.** *Firms choose policies that only depend on immediate past actions and the current state. That is, for each $t \geq 1$, $\sigma_t^i \equiv \sigma_t^i(p|\boldsymbol{p}_{t-1},s_t)$ is a function of only $p \in \mathcal{A}$, $\boldsymbol{p}_{t-1} \in \mathcal{A}^n$ and $s_t \in \mathcal{S}$ and is independent of the time.*

We remark that while we use in different places the general term $\sigma_t^i(p|\boldsymbol{p}_{t-1},s_t)$, the above assumption implies that it equals $\sigma_1^i(p|\boldsymbol{p}_{t-1},s_t)$ for all $t \geq 1$ and $\sigma_0^i(p|s_0)$ for $t=0$. Similarly, whenever we write $\boldsymbol{\sigma}$ we actually mean $\boldsymbol{\sigma}=(\boldsymbol{\sigma}_0,\boldsymbol{\sigma}_1)$.

Solution Concept. We study the one-memory subgame perfect Nash equilibrium (SPE) of the stochastic game. We clarify this definition in Section 2.2.3 and establish the existence of an SPE under specific assumptions in Section 2.3.

Additional Notation: (i) We denote $M=|\mathcal{A}^n|$ and write the set $\mathcal{S}\times\mathcal{A}^n$ as follows

$$\mathcal{S}\times\mathcal{A}^n=\left\{(s^1,\boldsymbol{p}^1),\cdots,(s^1,\boldsymbol{p}^M),\cdots,(s^r,\boldsymbol{p}^1),\cdots,(s^r,\boldsymbol{p}^M)\right\}. \tag{2.3}$$

(ii) The set of policies available at time $t \geq 0$ for firm i is denoted by $\boldsymbol{\Sigma}_t^i$. Using the enumeration in (2.3) and the notation $\hat{M}=(m+1)rM$, the set $\boldsymbol{\Sigma}_t^i$ can be identified with

$$\begin{aligned}\boldsymbol{\Sigma}_t^i=\{(\sigma_t^i(a^0|\boldsymbol{p}^1,s^1),\ldots,\sigma_t^i(a^m|\boldsymbol{p}^1,s^1),\ldots,\sigma_t^i(a^0|\boldsymbol{p}^M,s^r),\ldots,\sigma_t^i(a^m|\boldsymbol{p}^M,s^r))\\ \in[0,1]^{\hat{M}}\text{ s.t. }\sum_{k=0}^{m}\sigma_t^i(a^k|\boldsymbol{p}_0,s_1)=1\quad\forall(s_1,\boldsymbol{p}_0)\in\mathcal{S}\times\mathcal{A}^n\}.\end{aligned} \tag{2.4}$$

It follows from (2.4) that $\boldsymbol{\Sigma}_t^i$ is an $\hat{M}-1$ simplex, and consequently it is a compact and convex subset of $\mathbb{R}^{(m+1)rM}$.

The set of policies at time $t \geq 0$ for all firms is $\boldsymbol{\Sigma}_t := \times_{i=1}^n \boldsymbol{\Sigma}_t^i$. The set of all policies is $\boldsymbol{\Sigma} := \times_{t\geq 0}\boldsymbol{\Sigma}_t$.

(iii) A policy profile for time $t \geq 0$ contains the policies for all firms at that time and is described by $\boldsymbol{\sigma}_t = (\sigma_t^i)_{i=1}^n$. The sequence $\boldsymbol{\sigma}_t^{-i} = (\sigma_t^j)_{j\neq i}$ is a policy profile that excludes only σ_t^i. Similarly, $\boldsymbol{\Sigma}_t^{-i} := \times_{j\neq i}\boldsymbol{\Sigma}_t^j$. For each $i \in [n]$, we interchange between $(\sigma_t^i, \boldsymbol{\sigma}_t^{-i})$ and $(\boldsymbol{\sigma}_t^j)_{j=1}^n$.

(iv) For $\boldsymbol{\sigma}_t \in \boldsymbol{\Sigma}_t$, $\boldsymbol{p}_{t-1} \in \mathcal{A}^n$, $s_t \in \mathcal{S}$ and a real-valued function $g : \mathcal{A}^n \times \mathcal{S} \to \mathbb{R}$:

$$\mathbb{E}_{\boldsymbol{\sigma}_t}\left[g(\boldsymbol{p}, s)|\boldsymbol{p}_{t-1}, s_t\right] := \sum_{\boldsymbol{p}_t \in \mathcal{A}^n} \sigma_t(\boldsymbol{p}_t|\boldsymbol{p}_{t-1}, s_t) g(\boldsymbol{p}_t, s_t). \tag{2.5}$$

(v) For $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_t)_{t\geq 0} \in \boldsymbol{\Sigma}$, $s_0 \in \mathcal{S}$, $\mathbb{P}$ defined in (2.1) and a family of real-valued functions, $g_t : \mathcal{A}^n \times \mathcal{S} \to \mathbb{R}$, $t \geq 0$, we define

$$\mathbb{E}_{\boldsymbol{\sigma},\mathbb{P}}\left[\sum_{t=0}^{\infty} g_t(\boldsymbol{p}_t, s_t)|s_0\right] := \lim_{T\to\infty} \mathbb{E}_{\boldsymbol{\sigma},\mathbb{P}}\left[\sum_{t=0}^{T} g_t(\boldsymbol{p}_t, s_t)|s_0\right],$$

whenever the limit exists, where for each $T \geq 1$,

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{\sigma},\mathbb{P}}\left[\sum_{t=0}^{T} g_t(\boldsymbol{p}_t, s_t)|s_0\right] \\ &= \sum_{\boldsymbol{p}_0 \in \mathcal{A}^n} \sigma_0(\boldsymbol{p}_0|s_0)\left\{g_0(\boldsymbol{p}_0, s_0) + \sum_{s_1 \in \mathcal{S}} \mathbb{P}(s_1|\boldsymbol{p}_0, s_0)\mathbb{E}_{(\boldsymbol{\sigma}_t)_{t\geq 1},\mathbb{P}}\left[\sum_{t=1}^{T} g_t(\boldsymbol{p}_t, s_t)|\boldsymbol{p}_0, s_1\right]\right\}\end{aligned} \tag{2.6}$$

and for each $1 \leq k \leq T-1$

$$\mathbb{E}_{(\boldsymbol{\sigma}_t)_{t\geq k},\mathbb{P}}\left[\sum_{t=k}^{T} g_t(\boldsymbol{p}_t, s_t)|\boldsymbol{p}_{k-1}, s_k\right] = \sum_{\boldsymbol{p}_k\in\mathcal{A}^n} \sigma_k(\boldsymbol{p}_k|\boldsymbol{p}_{k-1}, s_k) g_k(\boldsymbol{p}_k, s_k)$$
$$+ \sum_{\boldsymbol{p}_k\in\mathcal{A}^n} \sigma_k(\boldsymbol{p}_k|\boldsymbol{p}_{k-1}, s_k) \sum_{s_{k+1}\in\mathcal{S}} \mathbb{P}(s_{k+1}|\boldsymbol{p}_k, s_k)\mathbb{E}_{(\boldsymbol{\sigma}_t)_{t\geq k+1},\mathbb{P}}\left[\sum_{t=k+1}^{T} g_t(\boldsymbol{p}_t, s_t)|\boldsymbol{p}_k, s_{k+1}\right]. \tag{2.7}$$

2.2.1 The V^i-Functions

The initial state $s_0 \in \mathcal{S}$ along with a profile of policies for all firms $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ determine the evolution of the stochastic game via conditional value functions, which we clarify in this section. Let $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_t)_{t=0}^{\infty} \in \boldsymbol{\Sigma}$ be a one-memory policy. We recall that by Assumption 1, for each firm $i \in [n]$, $\boldsymbol{\sigma}$ is characterized by two policies: (i) $\sigma_0^i(\cdot|s_0)$ at $t = 0$; and (ii) $\sigma_1^i(\cdot|\boldsymbol{p}_{t-1}, s_t)$ at $t \geq 1$. We will thus obtain conditional value functions for $t = 0$ and $t = 1$.

We define the conditional value function using the definition of $\mathbb{E}_{\boldsymbol{\sigma},\mathbb{P}}$ in (2.6) and (2.7). We recall that $\mathbb{P}$ is the distribution defined in (2.1), and $\pi^i : \mathcal{A}^n \times \mathcal{S} \to \mathbb{R}$, $i \in [n]$, are the profit functions. Let $\delta_i \in (0, 1)$ denote the discount factor for firm $i \in [n]$, which represents the present value of future profits. For $\boldsymbol{\sigma} = (\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^{-i}) \in \boldsymbol{\Sigma}$ and $s_0 \in \mathcal{S}$, the conditional value function at time $t = 0$ of firm i is given by

$$\tilde{V}_0^i(s_0, \boldsymbol{\sigma}^i|\boldsymbol{\sigma}^{-i}) := \mathbb{E}_{\boldsymbol{\sigma},\mathbb{P}}\left[\sum_{t=0}^{\infty} \delta_i^t \pi^i(\boldsymbol{p}_t, s_t)\Big|s_0\right]. \tag{2.8}$$

Given state s_0 at time $t = 0$, (2.8) measures the expected payoff that firm i receives after playing the infinite stochastic game using $\boldsymbol{\sigma}^i$, while firms other than i follow $\boldsymbol{\sigma}^{-i}$. Since $\pi^i(\boldsymbol{p}_t, s_t)$ is bounded by $\sup_{(s,\boldsymbol{p})\in\mathcal{S}\times\mathcal{A}^n} |\pi^i(\boldsymbol{p}, s)|$ for all $i \in [n]$ and $t \geq 0$, (2.8) is bounded by $(1-\delta_i)^{-1} \sup_{(s,\boldsymbol{p})\in\mathcal{S}\times\mathcal{A}^n} |\pi^i(\boldsymbol{p}, s)|$ and thus well-defined.

Next, we characterize the conditional value function of firm i at time $t = 1$. For $s_1 \in \mathcal{S}$,

$\boldsymbol{p}_0 \in \mathcal{A}^n$ and $\boldsymbol{\sigma}_1 = (\sigma_1^i, \boldsymbol{\sigma}_1^{-i}) \in \boldsymbol{\Sigma}_1$, the conditional value function of firm i at time $t = 1$ is given by

$$\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i}) := \mathbb{E}_{\boldsymbol{\sigma}_1, \mathbb{P}} \left[\sum_{t=1}^{\infty} \delta_i^{t-1} \pi^i(\boldsymbol{p}_t, s_t) \Big| \boldsymbol{p}_0, s_1 \right]. \tag{2.9}$$

For the pair $(s_1, \boldsymbol{p}_0)$ at time $t = 1$, (2.9) measures the expected payoff that firm i receives after playing the infinite stochastic game using σ_1^i, while firms other than i follow $\boldsymbol{\sigma}_1^{-i}$. If firm i uses a policy $\sigma_1^i \in \boldsymbol{\Sigma}_1^i$ such that $\sigma_1^i(\tilde{a}|\boldsymbol{p}_0, s_1) = 1$ for $\tilde{a} \in \mathcal{A}$ and for all $(\boldsymbol{p}_0, s_1) \in \mathcal{A}^n \times \mathcal{S}$, we write $\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \tilde{a}|\boldsymbol{\sigma}_1^{-i})$ instead of $\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i|\boldsymbol{\sigma}_1^{-i})$.

For technical reasons that will be explained in the next section, it is useful to define a $\boldsymbol{V}_1$ vector function. Its definition below uses a vector $\boldsymbol{v}$ whose coordinates are indexed by $i \in [n]$ and $(s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n$. In view of the enumeration of $\mathcal{S} \times \mathcal{A}^n$ in (2.3), $\boldsymbol{v} \in \mathbb{R}^{nrM}$. The $\boldsymbol{V}_1$ vector function is given by

$$\boldsymbol{V}_1 : \boldsymbol{\Sigma}_1 \times \boldsymbol{\Sigma}_1 \times \mathbb{R}^{nrM} \longrightarrow \mathbb{R}^{nrM} \text{ s.t. } (\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, \boldsymbol{v}) \mapsto \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, \boldsymbol{v})$$

where the $(i, s_1, \boldsymbol{p}_0)$-coordinate of $\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, \boldsymbol{v})$ is given by

$$\begin{aligned} &\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, \boldsymbol{v})_{i, s_1, \boldsymbol{p}_0} \\ &:= \sum_{\boldsymbol{p}_1 \in \mathcal{A}^n} \tau_1^i(p_1^i | \boldsymbol{p}_0, s_1) \sigma_1^{-i}(\boldsymbol{p}_1^{-i} | \boldsymbol{p}_0, s_1) \left[\pi^i(\boldsymbol{p}_1, s_1) + \delta_i \sum_{s_2 \in \mathcal{S}} \mathbb{P}(s_2 | \boldsymbol{p}_1, s_1) v_{i, s_2, \boldsymbol{p}_1} \right]. \end{aligned} \tag{2.10}$$

For the pair $(s_1, \boldsymbol{p}_0)$, (2.10) measures firm's i expected payoff from time $t = 1$ to time $t = 2$, assuming firm i follows τ_1^i at time $t = 1$, firms other than i follow $\boldsymbol{\sigma}_1^{-i}$, and the payoffs for all firms at time $t = 2$ are given by the vector $\boldsymbol{v}$. Note that the $(i, s_1, \boldsymbol{p}_0)$-coordinate of $\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, \boldsymbol{v})$ only depends on τ_1^i, which is why we often write $\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})_{i, s_1, \boldsymbol{p}_0}$ instead of $\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, \boldsymbol{v})_{i, s_1, \boldsymbol{p}_0}$, when no confusion can arise.

2.2.2 Further Clarification of $\tilde{V}_1^i$ and its Relationship with $\boldsymbol{V}_1$

The following fundamental proposition formulates a Bellman Equation for $\tilde{V}_1^i$. We use it to interpret $\tilde{V}_1^i$ as a weighted sum of conditional expectations and to directly relate $\tilde{V}_1^i$ to $\boldsymbol{V}_1$.

Proposition 1 (Lemma 1 of Fink (1964))**.** *Let $i \in [n]$ and $\boldsymbol{\sigma}_1 = (\sigma_1^i, \boldsymbol{\sigma}_1^{-i}) \in \boldsymbol{\Sigma}_1$. For each $(s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n$, $\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i})$ satisfies the following Bellman Equation,*

$$\begin{aligned} &\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i}) \\ &= \sum_{\boldsymbol{p}_1 \in \mathcal{A}^n} \sigma_1(\boldsymbol{p}_1 | \boldsymbol{p}_0, s_1) \left[\pi^i(\boldsymbol{p}_1, s_1) + \delta_i \sum_{s_2 \in \mathcal{S}} \mathbb{P}(s_2 | \boldsymbol{p}_1, s_1) \tilde{V}_1^i(s_2, \boldsymbol{p}_1, \sigma_1^i | \boldsymbol{\sigma}_1^{-i}) \right]. \end{aligned} \tag{2.11}$$

Moreover, the system of equations given by (2.11), which has rM equations and rM variables, $\{\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i})\}_{(s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n}$, has a unique solution.

We remark that this observation provides a simpler technical understanding of the conditional value function at time $t = 1$. Its original definition in (2.9) is an expectation of an infinite series, but the proof of Proposition 1 implies that it is an expectation of a simpler and easily interpretable finite sum. Indeed, following this proof in Appendix Appendix A.10.1, one can notice that for each $(s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n$, $\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i})$ is a weighted sum of the entries of $\mathbb{E}_{\boldsymbol{\sigma}_1}[\pi^i] := (\mathbb{E}_{\boldsymbol{\sigma}_1}[\pi^i | \boldsymbol{p}^1, s^1], \cdots, \mathbb{E}_{\boldsymbol{\sigma}_1}[\pi^i | \boldsymbol{p}^M, s^r])^T \in \mathbb{R}^{rM}$. Moreover, such weights are uniquely determined by the policies in $\boldsymbol{\sigma}_1$ and the transition probability $\mathbb{P}$ (see (A.42) in Appendix Appendix A.10.1).

Equation (2.11) also establishes a relationship between the conditional value function from time $t = 1$ and the $\boldsymbol{V}_1$ vector function. It implies that

$$\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i}) = \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_1, \tilde{\boldsymbol{v}})_{i, s_1, \boldsymbol{p}_0}, \tag{2.12}$$

for each $(i, s_1, \boldsymbol{p}_0)$-coordinate, where $\tilde{\boldsymbol{v}}_{i, s_1, \boldsymbol{p}_0} := \tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i})$. To see this, compare (2.11) with (2.10) written with $\tau_1 = \sigma_1$ and $\boldsymbol{v} = \tilde{\boldsymbol{v}}$.

2.2.3 Nash equilibrium

Using the definitions of the two conditional value functions at times $t = 0$ and $t = 1$, we define the concepts of a Nash equilibrium from time $t = 1$ and a subgame perfect Nash equilibrium.

A policy $\sigma_1^{i*} \in \mathbf{\Sigma}_1^i$ is called a best-response policy to $\boldsymbol{\sigma}_1^{-i} \in \mathbf{\Sigma}_1^{-i}$ if for all $(s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n$,

$$\sigma_1^{i*} \in \text{argmax}_{\sigma_1^i \in \mathbf{\Sigma}_1^i} \tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i}), \tag{2.13}$$

where $\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i})$ is given by (2.9). We say that $\boldsymbol{\sigma}_1^* \in \mathbf{\Sigma}_1$ is a Nash equilibrium from time $t = 1$, if for all $i \in [n]$, σ_1^{i*} is a best-response policy to $\boldsymbol{\sigma}_1^{-i*}$. In other words, $\boldsymbol{\sigma}_1^* \in \mathbf{\Sigma}_1$ is a Nash equilibrium from time $t = 1$, if for all $i \in [n]$, and $(s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n$,

$$\sigma_1^{i*} \in \text{argmax}_{\sigma_1^i \in \mathbf{\Sigma}_1^i} \tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i*}). \tag{2.14}$$

We define a subgame perfect Nash equilibrium as a profile $(\boldsymbol{\sigma}_0^*, \boldsymbol{\sigma}_1^*)$ such that $\boldsymbol{\sigma}_1^*$ is a Nash equilibrium from time $t = 1$ and $\boldsymbol{\sigma}_0^* \in \mathbf{\Sigma}_0$ satisfies for each $i \in [n]$

$$\sigma_0^{i*} \in \text{argmax}_{\sigma_0^i \in \mathbf{\Sigma}_0} \tilde{V}_0^i(s_0, (\sigma_0^i, \sigma_1^{i*}) | (\boldsymbol{\sigma}_0^{-i*}, \boldsymbol{\sigma}_1^{-i*})). \tag{2.15}$$

2.3 Existence of One-Memory SPEs

We establish the existence of a one-memory SPE and formulate an algorithm for checking whether a given profile is a one-memory SPE. The existence result is developed in three stages. First, Theorem 1, which is Theorem 2 of Fink (1964), shows that, starting from time $t = 1$, the $\boldsymbol{V}_1$ vector function possess stationary points. It also provides useful properties of these stationary points. Next, Theorem 2 shows that the latter stationary points are a Nash equilibrium from time $t = 1$. Lastly, Theorem 3 shows the existence of a one-memory SPE for the stochastic game. We demonstrate the application of this theory to grim trigger strategies in Section 2.3.1.

Theorem 1 (Existence of stationary points with special properties (Fink, 1964))**.** *There exist* $\boldsymbol{\sigma}_1^* \in \boldsymbol{\Sigma}_1$ *and* $\boldsymbol{v}^* \in \mathbb{R}^{nrM}$ *satisfying*

$$\boldsymbol{v}^* = \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \boldsymbol{\sigma}_1^*, \boldsymbol{v}^*) \tag{2.16}$$

and

$$\boldsymbol{v}^*_{i,s_1,\boldsymbol{p}_0} = \max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \sigma_1^i, \boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0} \;\; \forall\, (i, s_1, \boldsymbol{p}_0) \in [n] \times \mathcal{S} \times \mathcal{A}^n. \tag{2.17}$$

Theorem 2 (Existence of Nash Equilibrium from time $t = 1$)**.** *Suppose that* $\boldsymbol{\sigma}_1^* \in \boldsymbol{\Sigma}_1$ *and* $\boldsymbol{v}^* \in \mathbb{R}^{nrM}$ *satisfy* (2.16) *and* (2.17)*. Then, for each* $i \in [n]$ *and* $(s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n$*,*

$$\max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \sigma_1^i, \boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0} = \max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i*}). \tag{2.18}$$

Moreover, $\boldsymbol{\sigma}_1^*$ *is a Nash equilibrium from time* $t = 1$*.*

Theorem 3 (Existence of the one-memory SPE)**.** *If* $\boldsymbol{\sigma}_1^* \in \boldsymbol{\Sigma}_1$ *is a Nash equilibrium from time* $t = 1$*, then there exists* $\boldsymbol{\sigma}_0^* \in \boldsymbol{\Sigma}_0$ *such that* $\boldsymbol{\sigma}^* = (\boldsymbol{\sigma}_0^*, \boldsymbol{\sigma}_1^*)$ *is a one-memory SPE of the stochastic game.*

This theory suggests the following three-step algorithm for proving that a given profile is a one-memory SPE. If one can only verify the first two steps of the algorithm, then the given profile is a Nash equilibrium from time $t = 1$. We frequently use this algorithm in our proofs.

Algorithm 1 (Proving that a given profile is a one-memory SPE)**.** *Let* $(\boldsymbol{\sigma}_0^g, \boldsymbol{\sigma}_1^g)$ *be a given one-memory strategy profile. The following algorithm guides the proof that this profile is* *an SPE. Its first two steps are used for proving a Nash equilibrium from time $t = 1$.*

1. *Plug in $\boldsymbol{\sigma}_1^g$ into equation (2.16) and solve it as a linear system with unknowns $v_{i,s_1,\boldsymbol{p}_0}^g$ for each $(i, s_1, \boldsymbol{p}_0)$-coordinate.*

2. *Plug in $\boldsymbol{v}^g$ and $\boldsymbol{\sigma}_1^g$ into (2.17) and show that $\boldsymbol{v}^g$ is a fixed point of the operator $v_{i,s_1,\boldsymbol{p}_0} \mapsto \max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{\sigma}_1^g, \sigma_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}$.*

3. *Show that $\boldsymbol{\sigma}_0^g$ satisfies (2.15).*

Comments on the Proofs of Theorems 1, 2 and 3. The proof of Theorem 1 is due to Fink (1964). For completeness, Appendix Appendix A.10 rewrites Fink's proof using our notation, while including many of the missing details in Fink (1964). We find it necessary to refer to the rewritten proof when establishing the theories of Sections 2.3.1 and 2.4.

We note that Theorem 1 itself is not enough to guarantee the existence of a Nash equilibrium from time $t = 1$. To obtain such a guarantee one needs to prove (2.18) and then use Theorem 1. In general, the equality $\boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \sigma_1^i, \boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0} = \tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i*})$ need not be true for all $\sigma_1^i \in \boldsymbol{\Sigma}_1^i$ and should not be confused with (2.12). The proof that equation (2.18) holds true requires some detailed inequalities (see Appendix Appendix A.1).

To prove Theorem 3, we show that finding a solution for (2.15) is equivalent to finding a static Nash equilibrium in mixed strategies of a particular finite game. We recall that an n-person finite game is any set $\{(X^i, q^i)\}_{i=1}^n$ where X^i is a nonempty finite set of actions and $q^i : X := \times_{i=1}^n X^i \to \mathbb{R}$ is the profit for the ith player. A mixed strategy for agent i is a probability mass function γ^i on X^i. Given $\boldsymbol{\gamma} = (\gamma^i)_{i=1}^n$, the expected return for agent i is given by $\mathbb{E}_{\boldsymbol{\gamma}} q^i := \sum_{\boldsymbol{x} \in X} \boldsymbol{\gamma}(\boldsymbol{x}) q^i(\boldsymbol{x})$, where $\boldsymbol{\gamma}(\boldsymbol{x})$ denotes the product of $\gamma^i(x^i)$ for $i \in \{1, \dots, n\}$. From a theorem by Nash (1950), any n-person finite game has a Nash equilibrium in mixed strategies.

For each $(\boldsymbol{p}_0, s_0) \in \mathcal{A}^n \times \mathcal{S}$, we define the following quantity

$$Q^{i*}(\boldsymbol{p}_0, s_0) := \pi^i(\boldsymbol{p}_0, s_0) + \delta_i \sum_{s_1 \in \mathcal{S}} \mathbb{P}(s_1|\boldsymbol{p}_0, s_0) v^*_{i,s_1,\boldsymbol{p}_0}. \tag{2.19}$$

This quantity appeared in Hu and Wellman (2003) and was referred to as the Nash Q-function of firm i at $(\boldsymbol{p}_0, s_0)$. We show (see Appendix Appendix A.2) that

$$\tilde{V}^i_0(s_0, \boldsymbol{\sigma}^i|\boldsymbol{\sigma}^{-i}) = \mathbb{E}_{\boldsymbol{\sigma}_0}[Q^{i*}(\boldsymbol{p}, s)|s_0]. \tag{2.20}$$

In view of this equation and the use of expected return in an n-person finite game, finding $\boldsymbol{\sigma}^*_0 \in \boldsymbol{\Sigma}_0$ satisfying (2.15) for each $i \in [n]$ is equivalent to finding a Nash equilibrium of the finite game $\{(\mathcal{A}, Q^{i*})\}^n_{i=1}$, where $\mathcal{A}$ is the set of actions from Section 2.2.

2.3.1 Application: Grim Trigger Strategies as an SPE

The theory of Section 2.3 applies to a large family of stochastic games. Due to this generality, we can prove some non-trivial statements about how collusion can be sustained with one-memory policies. In particular, we provide sufficient conditions under which grim trigger strategies supporting a collusive-enabling price are an SPE of the stochastic game.

First, we specify these sufficient conditions. The first assumption aligns our stochastic game with the key aspect of the dynamic Bertrand competition model, which is the existence of a Nash equilibrium price and a collusive-enabling price.

Assumption 2. *There exists a Nash equilibrium price $\mathbf{p}^* = (p^*, \dots, p^*) \in \mathcal{A}^n$ of the one-stage game $\{(\mathcal{A}, \pi^i)\}^n_{i=1}$, where π^i is given by (2.2). There exists a price $\mathbf{p}^C = (p^C, \dots, p^C)$ such that $\pi^i(\boldsymbol{p}^*) < \pi^i(\boldsymbol{p}^C)$ for each $i \in [n]$. We refer to $\boldsymbol{p}^*$ and $\mathbf{p}^C$ as the competition and collusive-enabling prices, respectively.*[4]

[4]This assumption is satisfied by traditional Bertrand competition models (see, e.g., Tirole (1988)), but it is also satisfied by more recent models such as those of platform competition in two-sided markets (see, e.g., Dewenter et al. (2011) and Chica et al. (2023a)).

We further assume two simple conditions that we only use in this section.

Assumption 3. *We require the following two conditions:*

(i) $|\mathcal{S}| = 1$.

(ii) For each $i \in [n]$ *and* $\pi^{m,i} := \max_{p^i \in \mathcal{A} \setminus \{p^C\}} \pi^i(p^i, (\boldsymbol{p}^C)^{-i})$, $\frac{\pi^{m,i} - \pi^i(\boldsymbol{p}^C)}{\pi^{m,i} - \pi^i(\boldsymbol{p}^*)} \leq \delta_i < 1$.

Condition (i) in Assumption 3 reduces our stochastic game to an infinite repeated game. Condition (ii) provides a lower bound on δ_i and thus in view of (2.8), it lower bounds the level of patience needed by the firms. The lower bound uses the quantity $\pi^{m,i}$, which is the best response payoff of firm i when other firms charge the collusive-enabling price p^C, where we note that by definition $\pi^{m,i} \geq \pi^i(\boldsymbol{p}^C)$. The bound is the ratio of the distance between $\pi^{m,i}$ and the collusive-enabling payoff, $\pi^i(\boldsymbol{p}^C)$, and the distance between $\pi^{m,i}$ and competition payoff $\pi^i(\boldsymbol{p}^*)$.

Next, we review the grim trigger strategy and formulate the main proposition of this section. The grim trigger strategy (Friedman, 1985) in our setting (assuming Assumption 2) is a policy in which a firm cooperates by choosing the price p^C as long as all other firms have chosen the price p^C in the previous stage. If, on the other hand, at least one firm did not cooperate in the previous stage and chose a price $p^i \neq p^C$, other firms will defect by playing p^* forever. We note that because $\boldsymbol{p}^*$ is a Nash equilibrium, firm i has incentives to also choose p^* going forward (this is verified for our setting in the proposition below). Therefore, firm i is punished since it will receive $\pi^i(\boldsymbol{p}^*)$ forever without having any competitive advantage (since everyone else is doing the same), whereas it previously received $\pi^i(\boldsymbol{p}^C)$. In our setting of one-memory stochastic games, the grim trigger strategy is defined as the following one-memory policy: $\boldsymbol{\sigma}^f = (\sigma_0^f, \sigma_1^f)$, where

$$\sigma_0^f(p^C) = 1, \sigma_1^f(p^C|\boldsymbol{p}^C) = 1 \text{ and } \forall \boldsymbol{p}_0 \in \mathcal{A}^n, \boldsymbol{p}_0 \neq \boldsymbol{p}^C, \sigma_1^f(p^*|\boldsymbol{p}_0) = 1.$$

Proposition 2 (The grim trigger strategy is a one-memory SPE)**.** *Under the assumptions of Section 2.2 and Assumptions 2 and 3, the grim trigger strategy is an SPE of the stochastic game. Moreover,*

$$\tilde{V}_0^i(\boldsymbol{\sigma}^f) = \frac{1}{1-\delta_i}\pi^i(\boldsymbol{p}^C). \tag{2.21}$$

The proof of Proposition 2 can be found in Appendix Appendix A.3; it mainly uses Algorithm 1. The fact that grim trigger strategies can be seen as an SPE is not new (see, e.g., Friedman (1985) and Osborne (1994)). However, our proof demonstrates a straightforward way of applying Algorithm 1 to determine an SPE, where the other proofs are long and are different from our underlying theory.

2.4 Q-Learning and Collusion

We establish important properties of Q-learning (Watkins and Dayan, 1992), which is the most common reinforcement learning algorithm. Section 2.4.1 presents a version of Q-learning with no experimentation, while adapting it to our stochastic setting of Section 2.2. Furthermore, it connects between the fixed points of the algorithm and the $\boldsymbol{V}^i$-functions of Section 2.2.1. Specifically, it shows that these fixed points capture the value of the stochastic game at time $t = 1$ for a specific type of strategies, named induced strategies. Moreover, it establishes sufficient conditions for these induced strategies to be a Nash equilibrium from time $t = 1$. We remark that the induced strategies are one-memory strategies and thus we can naturally apply the results obtained for the latter strategies in the previous section. Section 2.4.2 presents a version of Q-learning with bounded experimentation. It further provides sufficient conditions for the convergence of this algorithm in stochastic games satisfying Assumptions 2 and 3-(i), and in particular for the traditional dynamic Bertrand competition model. Moreover, it establishes sufficient conditions under which Q-learning firms learn that choosing supracompetitive prices is optimal. It also presents sufficient conditions under which these supracompetitive prices are supported by either naive collusion, grim trigger strategies or increasing strategies.

2.4.1 A Relationship of a Q-Learning Algorithm with the Stochastic Game

We formulate a version of the Q-learning algorithm with no experimentation, while assuming the multi-agent setting of Section 2.2. We then establish the relationship of the Q-function of this algorithm with the value functions, V_1^i and $\tilde{V}_1^i$, of the stochastic game. The basic idea of this algorithm is to find a policy that maximizes (2.9) given the policies of all other agents. The algorithm takes as input $Q_0^i : \mathcal{S} \times \mathcal{A}^{n+1} \to \mathbb{R}$ for $i \in [n]$, as well as several parameters, and output $Q_t^i : \mathcal{S} \times \mathcal{A}^{n+1} \to \mathbb{R}$ for $i \in [n]$ and $t \geq 1$. We use the notation $\boldsymbol{s} = (s, \boldsymbol{p}) \in \mathcal{S} \times \mathcal{A}^n$.

Algorithm 2 (Q-learning with no experimentation)**.** *Arbitrarily fix $\boldsymbol{p}_0 \in \mathcal{A}^n$ and $s_1 \in \mathcal{S}$. For each $(\boldsymbol{s}, p) \in \mathcal{S} \times \mathcal{A}^{n+1}$ and $j \in [n]$, let $Q_0^j(\boldsymbol{s}, p) = 0$. At time $t \geq 1$, firm i observes $\boldsymbol{s}_t = (s_t, \boldsymbol{p}_{t-1}) \in \mathcal{S} \times \mathcal{A}^n$ and updates its Q-values using the following rule, for each $(\boldsymbol{s}, p) \in \mathcal{S} \times \mathcal{A}^{n+1}$,*

$$Q_{t+1}^i(\boldsymbol{s}, p) = (1 - \alpha_t) Q_t^i(\boldsymbol{s}, p) + \alpha_t \left\{ \pi^i(\boldsymbol{p}_t, s) + \delta_i \mathbb{E}_{\boldsymbol{s}_{t+1}} \left[\max_{a \in \mathcal{A}} Q_t^i(\boldsymbol{s}_{t+1}, a) \right] \right\}, \tag{2.22}$$

where both the profit function $\pi^i(\boldsymbol{p}_t, \boldsymbol{s})$ and rates $\alpha_t = \alpha_t(\boldsymbol{s}, p) \in [0, 1]$ for $t \geq 1$ are parametric choices of the algorithm. For $t \geq 1$, $\alpha_t = 0$ for each $(\boldsymbol{s}, p) \neq (\boldsymbol{s}_t, p_t^i)$, That is α_t is only considered for the states and prices visited at time t, $(\boldsymbol{s}_t, p_t^i)$. Then, with uniform probability, firm i chooses a price among

$$p_t^i \in \operatorname{argmax}_{a \in \mathcal{A}} Q_t^i(\boldsymbol{s}_t, a). \tag{2.23}$$

Firm i then observes both prices $\boldsymbol{p}_t$ and profits $(\pi^j(\boldsymbol{p}_t, s_t))_{j=1}^n$, and randomly draws $\boldsymbol{s}_{t+1} = (s_{t+1}, \boldsymbol{p}_t)$ with probability $\mathbb{P}(s_{t+1}|\boldsymbol{p}_t, s_t)$, where $\mathbb{P}$ is another parametric choice of the algorithm.

Suppose that $\boldsymbol{Q}^* = (Q^{i*})_{i=1}^n$ is a fixed point of Algorithm 2 with constant learning rate $\alpha_t = \alpha \in (0,1]$ for each $t \geq 0$. Assume that starting from time $t = 1$, firms use Q^{i*} to play the stochastic game described in Section 2.2 as follows: Given $\boldsymbol{s} \in \mathcal{S} \times \mathcal{A}^n$, each firm $i \in [n]$ chooses

$$w^i(\boldsymbol{s}) \in \text{argmax}_{p \in \mathcal{A}} Q^{i*}(\boldsymbol{s}, p). \tag{2.24}$$

We denote $\boldsymbol{w}(\boldsymbol{s}) = (w^i(\boldsymbol{s}))_{i=1}^n$. The latter strategies are often referred to as the strategies induced by $\boldsymbol{Q}^*$. Moreover, $\boldsymbol{w}(\boldsymbol{s})$ is a one-memory strategy, since $\boldsymbol{s}$ depends on the price choices of the previous stage. The following proposition shows that if agents play the stochastic game following the strategies induced by $\boldsymbol{Q}^*$, then the *conditional* value function of firm i at time $t = 1$ (see (2.9)) coincides with Q^{i*} at the induced strategies.

Proposition 3 (Q^{i*} captures the value of the game at time $t = 1$)**.** *Assume $\alpha_t = \alpha \in (0,1]$ for each $t \geq 0$ and $(Q^{i*})_{i=1}^n$ is a fixed point of Algorithm 2. Then, for each $i \in [n]$ and $\boldsymbol{s} = (s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n$,*

$$Q^{i*}(\boldsymbol{s}, w^i(\boldsymbol{s})) = \tilde{V}_1^i(\boldsymbol{s}, w^i(\boldsymbol{s})|\boldsymbol{w}^{-i}(\boldsymbol{s})). \tag{2.25}$$

Note, however, that the latter property is not enough to show that the induced strategies are a Nash equilibrium from time $t = 1$. The following proposition shows a sufficient condition for the induced strategy to be a Nash equilibrium from time $t = 1$.

Proposition 4 (Sufficient condition for $\boldsymbol{Q}^*$ to induce a Nash equilibrium from time $t = 1$)**.** *Assume $\alpha_t = \alpha \in (0,1]$ for each $t \geq 0$, $\boldsymbol{Q}^*$ is a fixed point of Algorithm 2, and for each*

$i \in [n]$ and $\boldsymbol{s} = (s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n$,

$$w^i(\boldsymbol{s}) \in \text{argmax}_{p_1^i \in \mathcal{A}} \boldsymbol{V}_1(\boldsymbol{w}, p_1^i, \boldsymbol{Q}^*)_{i,\boldsymbol{s}}, \tag{2.26}$$

where $\boldsymbol{w} = \{w^i(\boldsymbol{s}) | i \in [n], \boldsymbol{s} \in \mathcal{S} \times \mathcal{A}^n\}$ and $\boldsymbol{V}_1$ is given by (2.10). Then, the strategy induced by $\boldsymbol{Q}^*$ is a Nash equilibrium from time $t = 1$.

Suppose that given a state $\boldsymbol{s} \in \mathcal{S} \times \mathcal{A}^n$, firms play a one-stage game with payoffs given by $(\boldsymbol{V}_1(\cdot, \cdot, \boldsymbol{Q}^*)_{i,\boldsymbol{s}})_{i \in [n]}$. In this case, Proposition 4 implies that if the induced strategy by $\boldsymbol{Q}^*$ is a Nash equilibrium of the latter one-stage game, then this strategy is a Nash equilibrium from time $t = 1$ for the stochastic game of Section 2.3. This observation is interesting since Algorithm 1 requires checking two conditions in order to decide whether a given profile is a Nash equilibrium from time $t = 1$. However, in the current case only one condition is needed because $\boldsymbol{w}(\boldsymbol{s})$ is induced from a fixed-point of Algorithm 2.

2.4.2 The Rise of Supracompetitive Prices and Collusion with the Q-Learning Algorithm

In this section, we study a version of Q-learning with bounded experimentation in stochastic games satisfying Assumptions 2 and 3-(i). We establish conditions for its convergence, formulate sufficient conditions for firms to learn that choosing supracompetitive prices is optimal, and demonstrate sufficient conditions under which these supracompetitive prices are supported by either naive collusion, grim trigger strategies or increasing strategies (defined below).

We will only use here Assumptions 2 and 3-(i) from Section 2.3.1. The first assumption ensures the presence of both a Nash equilibrium price and a price that facilitates collusion.

Assumption 3-(i) implies that states used in Algorithm 3 have the following form:[5]

$$\text{For } \ t \geq 1, \ \boldsymbol{s}_t = \boldsymbol{p}_{t-1} \in \mathcal{A}^n, \ \text{where } \boldsymbol{p}_{t-1} \text{ is the price choice at time } t-1.$$

Next, we define Q-learning with bounded experimentation. For this purpose we first define Q-learning with softmax, that is a version which replaces the maximal value of the Q-function with its soft-max value and thus allows "experimentation" with different prices. The desired Q-learning with bounded experimentation combines Algorithm 2 with Q-learning with softmax.

Algorithm 3 (Q-learning with softmax)**.** *This algorithm is similar to Algorithm 2, but instead of following (2.23) firm i chooses a price p_t^i by random draw according to the soft-max probability*

$$\sigma^i(p_t^i = a|\boldsymbol{s}_t) = \frac{e^{Q_t^i(\boldsymbol{s}_t,a)/\beta_t}}{\sum_{\tilde{a}\in\mathcal{A}} e^{Q_t^i(\boldsymbol{s}_t,\tilde{a})/\beta_t}}, \tag{2.27}$$

where $\beta_t > 0$ is chosen such that $\beta_t \to 0$ as $t \to \infty$, for example, $\beta_t = 1/t$.[6]

Algorithm 4 (Q-learning with bounded experimentation)**.** *Let $T > 0$ be an input parameter characterizing the size of experimentation. Firms follow Algorithm 3 from $t = 0$ to $t = T-1$ and follow Algorithm 2 from $t = T$ onward.*

Finally, we make the following assumption on the learning parameter α_t first described in Algorithm 3.

Assumption 4. *The learning rate α_t satisfies the following: (i) $0 < \alpha_t < 1$ for each $t \geq 0$ and $\sum_{t=T}^{\infty} \alpha_t = \infty$; (ii) for the fixed discount rate for firm i, $\delta_i \in (0,1)$, the following limit*

[5]We remark that this state choice has been a standard assumption in recent articles on algorithmic price discrimination (see, e.g. Calvano et al. (2020b), Klein (2021) and Chica et al. (2024a).)

[6]The no experimentation case can be recovered from (2.27) by replacing β_t with ϵ and letting $\epsilon \to 0$. In this case, $\mathbb{P}(p_t^i = \tilde{a}|\boldsymbol{s}) \to 1/|\text{argmax}_{a\in\mathcal{A}} Q_t^i(\boldsymbol{s},a)|$ if $\tilde{a} \in \text{argmax}_{a\in\mathcal{A}} Q_t^i(\boldsymbol{s},a)$; otherwise $\mathbb{P}(p_t^i = \tilde{a}|\boldsymbol{s}) \to 0$.

exists and satisfies

$$\alpha(\delta_i) := \lim_{t\to\infty} \sum_{k=T+1}^{t} \prod_{l=k+1}^{t} (1-\alpha_l(1-\delta_i))\alpha_k \in (0,\infty).$$

Part (i) in the above assumption is part of a standard assumption on the learning rates used by Watkins and Dayan (1992) to prove convergence of the Q-learning algorithm for single-agent models. Part (ii) ensures the convergence of the Q-learning algorithm in our stochastic setup.

The main result in this section is formulated as follows.

Theorem 4 (Q-learning convergence to supracompetitive prices)**.** *Suppose that Assumptions 2, 3-(i) and 4 hold, firms play with Algorithm 4 in the stochastic setting of Section 2.2, and for each $i \in [n]$, $p \in \mathcal{A} \setminus \{p^C\}$ and $\boldsymbol{s} \in \{\boldsymbol{p}_{T-1}, \boldsymbol{p}^C\}$:*

(i) $Q^i_T(\boldsymbol{s}, p^C) > Q^i_T(\boldsymbol{s}, p)$;

(ii) $\pi^i(\boldsymbol{p}^C) \geq (1-\delta_i) Q^i_T(\boldsymbol{p}^C, p)$.

Then, for any $\boldsymbol{p}_0 \in \mathcal{A}^n$ and each $t \geq T$, $p^i_t = p^C$. Moreover,

$$Q^{i*}_{\epsilon\to 0}(\boldsymbol{s}, p) := \lim_{t\to\infty} Q^i_t(\boldsymbol{s}, p) = \tag{2.28}$$
$$\begin{cases} \alpha(\delta_i)\pi^i(\boldsymbol{p}^C) & \text{if } (\boldsymbol{s}, p) = (\boldsymbol{p}^C, p^C), \\ (1-\alpha_T) Q^i_T(\boldsymbol{p}_{T-1}, p^C) + \alpha_T \left[\pi^i(\boldsymbol{p}^C) + \delta_i Q^i_T(\boldsymbol{p}^C, p^C)\right] & \text{if } (\boldsymbol{s}, p) = (\boldsymbol{p}_{T-1}, p^C) \text{ and } \boldsymbol{p}_{T-1} \neq \boldsymbol{p}^C, \\ Q^i_T(\boldsymbol{s}, p) & \text{otherwise.} \end{cases}$$

The proof of Theorem 4 can be found in Appendix Appendix A.4. Its basic idea is as follows. First, Algorithm 4 along with with condition (i) in Theorem 4 guarantee that the Q-learning algorithms choose $p^i_T = p^C$ for each $i \in [n]$ and $\boldsymbol{p}_0 \in \mathcal{A}^n$. Then, condition (ii) in Theorem 4 ensures that firms continue choosing $p^i_{T+1} = p^C$ in time $T+1$. Moreover, Assumption 4 ensures convergence in equation (2.28).

In order to discuss the relevance of Theorem 4, we recall the two key questions of our study: (i) What are sufficient conditions for firms to learn that choosing supracompetitive prices is optimal in the long run? (ii) Are these supracompetitive prices the result of punishment and reward strategies? Theorem 4 answers the first question and sheds light on the second question. First, it demonstrates sufficient conditions for Q-learning driven firms to consistently choose the collusive-enabling price p^C in every stage of the stochastic game. Thus showing that firms learn to choose supracompetitive prices in the long run. This result can help explaining the numerical results found by recent studies (e.g., Calvano et al. (2020b) and Chica et al. (2024a)), where authors numerically show that algorithms consistently learn to choose supracompetitive prices. Second, Theorem 4 characterizes the limiting Q-function $(Q^{i*}_{\epsilon\to 0})_{i=1}^n$. This observation together with Propositions 5, 6 and 7 below answer question (ii)

The rest of the section completes the answer to question (ii) described above. We first formulate the following proposition studying "naive collusion", that is, collusion without any punishment and reward behavior. It uses the notation $\boldsymbol{w}^*_{\epsilon\to 0} = (w^{i*}_{\epsilon\to 0})_{i=1}^n$ for the strategy induced by $(Q^{i*}_{\epsilon\to 0})_{i=1}^n$ defined in (2.28) (see (2.24) for the definition of induced strategies).

Proposition 5 (Naive Collusion)**.** *Suppose that Assumptions 2, 3-(i) and 4-(ii) hold, and* $\alpha(\delta_i)$ *satisfies* $\alpha(\delta_i)(1-\delta_i) > 1$ *for each* $i \in [n]$*. Furthermore, firms play with the induced strategies* $\boldsymbol{w}^*_{\epsilon\to 0}$ *in the stochastic setting of Section 2.2, and for each* $i \in [n]$ *and* $p \in \mathcal{A} \setminus \{p^C\}$

(i) $Q^i_T(\boldsymbol{s}, p^C) > Q^i_T(\boldsymbol{s}, p)$ *for each* $\boldsymbol{s} \in \mathcal{A}^n$*;*

(ii) $\pi^i(\boldsymbol{p}^C) \geq Q^i_T(\boldsymbol{p}_{T-1}, p) - \delta_i Q^i_T(\boldsymbol{p}^C, p)$ *for each* $\boldsymbol{s} \in \{\boldsymbol{p}_{T-1}, \boldsymbol{p}^C\}$.

Then, for each $\boldsymbol{s} \in \mathcal{A}^n$*,*

$$\boldsymbol{w}^*_{\epsilon\to 0}(\boldsymbol{s}) = \boldsymbol{p}^C.$$

Moreover, $\boldsymbol{w}^*_{\epsilon\to 0}$ is a Nash equilibrium from time $t = 1$ if and only if $\boldsymbol{p}^C$ is a Nash equilibrium of the one-stage game $(\pi^i(\cdot))_{i=1}^n$.

Proposition 5 shows sufficient conditions under which the strategies induced by the limiting Q-function $(Q^{i*}_{\epsilon\to 0})_{i=1}^n$ never display punishment and reward behavior. Indeed, there is no mechanism to punish a firm that chooses a price different than p^C. Instead, firms naively play by always choosing the collusive-enabling price. Therefore, this proposition implies that supracompetitive prices are not always the result of punishment and reward behavior. The final statement of Proposition 5 implies that unless $\boldsymbol{p}^C$ is a Nash equilibrium of the one-stage game $(\pi^i(\cdot))_{i=1}^n$, $\boldsymbol{w}^*_{\epsilon\to 0}$ cannot be a Nash equilibrium from time $t = 1$. However, in general, $\boldsymbol{p}^C$ is not a Nash equilibrium in most models of interest, such as traditional Bertrand competition or platform competition in two sided markets (see, e.g., Tirole (1988), Dewenter et al. (2011) and Chica et al. (2023a)). Finally, it is worth mentioning that Assumptions (i) and (ii) in Proposition 5 imply (i) and (ii) in Theorem 4. This is intuitive because achieving supracompetitive prices by naively choosing $\boldsymbol{p}^C$ in all states is a more difficult condition to meet than just achieving supracompetitive prices.

The following proposition shows sufficient conditions under which the strategies induced by $(Q^{i*}_{\epsilon\to 0})_{i=1}^n$ display punishment and reward behavior in a grim trigger fashion.

Proposition 6 (Grim Trigger Collusion)**.** *Suppose that Assumptions 2, 3-(i) and 4-(ii) hold, and $\alpha(\delta_i)$ satisfies $\alpha(\delta_i)(1-\delta_i) > 1$ for each $i \in [n]$. Furthermore, firms play with the induced strategies $\boldsymbol{w}^*_{\epsilon\to 0}$ in the stochastic setting of Section 2.2, and for each $i \in [n]$*

(i) $Q^i_T(\boldsymbol{s}, p^*) > Q^i_T(\boldsymbol{s}, p)$ *and* $Q^i_T(\boldsymbol{p}_{T-1}, p^*) > Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p)$ *for each* $\boldsymbol{s} \in \mathcal{A}^n \setminus \{\boldsymbol{p}^C, \boldsymbol{p}_{T-1}\}$, $p \in \mathcal{A} \setminus \{p^*\}$*;*

(ii) $\pi^i(\boldsymbol{p}^C) \geq (1-\delta_i) Q^i_T(\boldsymbol{p}^C, p)$ *for each* $p \in \mathcal{A} \setminus \{p^C\}$.

Then,

$$\boldsymbol{w}^*_{\epsilon\to 0}(\boldsymbol{s}) = \begin{cases} \boldsymbol{p}^C & \boldsymbol{s} = \boldsymbol{p}^C, \\ \boldsymbol{p}^* & \boldsymbol{s} \neq \boldsymbol{p}^C. \end{cases} \tag{2.29}$$

Moreover, under Assumption 3-(ii), $\boldsymbol{w}^*_{\epsilon\to 0}$ is a Nash equilibrium from time $t = 1$.

Proposition 6 shows sufficient conditions under which the strategies induced by the limiting Q-function $(Q^{i*}_{\epsilon\to 0})_{i=1}^n$ are equal to the grim trigger strategies starting from time $t = 1$ (see Section 2.3.1). By definition, these strategies display punishment and reward behavior since firms always collude, i.e., choose p^C, if all firms have chosen p^C in the previous stage. Otherwise, firms will defect by playing p^* forever. Recall that by Assumption 2, $\pi^i(\boldsymbol{p}^C) > \pi^i(\boldsymbol{p}^*)$. Thus, firms are better off by repeatedly choosing p^C in every stage of the game. Finally, it is worth mentioning that Assumptions (i) and (ii) in Proposition 6 are not in contradiction with (i) and (ii) in Theorem 4 (to see this note that in Theorem 4 the conditions are restricted to only two states: $\boldsymbol{s} \in \{\boldsymbol{p}_{T-1}, \boldsymbol{p}^C\}$). Therefore, in view of both Theorem 4 and Proposition 6, Q-learning firms may learn grim trigger strategies.

Now, punishment and reward schemes go beyond grim trigger strategies. In fact, Calvano et al. (2020b), Klein (2021) and Chica et al. (2024a) numerically show that algorithms can learn more elaborated collusive behavior, such as firms learning to progressively raise prices until reaching the collusive-enabling price $\boldsymbol{p}^C$ while using $\boldsymbol{p}^*$ as a threat in case of unilateral deviations. Proposition 7 below shows sufficient conditions under which the strategies induced by $(Q^{i*}_{\epsilon\to 0})_{i=1}^n$ mimic the latter behavior. It uses the following assumption.

Assumption 5. *There is a sequence of prices $\{p^l\}_{l=0}^{k+1} \subseteq \mathcal{A}$, where $p^l < p^{l+1}$ for each $l \in [k]$ and $(p_0, p^{k+1}) = (p^*, p^C)$, and denote $\boldsymbol{p}^l = (p^l)_{i=1}^n$. Furthermore, $\boldsymbol{p}_{T-1} \notin \{p^l\}_{l=0}^{k+1}$ and for each $i \in [n]$*

(i) $Q^i_T(\boldsymbol{p}^l, p^{l+1}) > Q^i_T(\boldsymbol{p}^l, p)$ *for each* $l \in [k]$, $p \in \mathcal{A} \setminus \{p^{l+1}\}$*;*

(ii) $Q_T^i(\boldsymbol{s}, p^*) > \max\{Q_T^i(\boldsymbol{s}, p), Q_{\epsilon\to 0}^i(\boldsymbol{p}_{T-1}, p^C)\}$ *for each* $p \in \mathcal{A} \setminus \{p^*\}$ *and* $\boldsymbol{s} \in \mathcal{A} \setminus \{p^l\}_{l=0}^{k+1}$ *with* $(\boldsymbol{s}, p) \neq (\boldsymbol{p}_{T-1}, p^C)$.

Proposition 7 (Increasing Strategies)**.** *Suppose that Assumptions 2, 3-(i), 4-(ii) and 5 hold, and* $\alpha(\delta_i)$ *satisfies* $\alpha(\delta_i)(1-\delta_i) > 1$ *for each* $i \in [n]$*. Furthermore, firms play with the induced strategies* $\boldsymbol{w}_{\epsilon\to 0}^*$ *in the stochastic setting of Section 2.2, and*

$$\pi^i(\boldsymbol{p}^C) \geq (1-\delta_i) Q_T^i(\boldsymbol{p}^C, p) \text{ for each } i \in [n] \text{ and } p \in \mathcal{A} \setminus \{p^C\}.$$

Then, for each $l \in [k]$

$$\boldsymbol{w}_{\epsilon\to 0}^*(\boldsymbol{s}) = \begin{cases} \boldsymbol{p}^C & \boldsymbol{s} = \boldsymbol{p}^C, \\ \boldsymbol{p}^{l+1} & \boldsymbol{s} = \boldsymbol{p}^l, \\ \boldsymbol{p}^* & \boldsymbol{s} \notin \{\boldsymbol{p}^l\}_{l=0}^{k+1}. \end{cases} \tag{2.30}$$

Proposition 7 shows sufficient conditions under which the strategies induced by the limiting Q-function $(Q_{\epsilon\to 0}^{i*})_{i=1}^n$ display an increasing behavior towards the collusive-enabling price p^C. Suppose that firms start at the Nash equilibrium price $\boldsymbol{p}^*$, following (2.30), firms will choose $\boldsymbol{p}^1$ in the next stage, and progressively increase their prices until reaching $\boldsymbol{p}^{k+1} = \boldsymbol{p}^C$. After any unilateral deviation, firms go back to the Nash equilibrium price and the increasing pattern follows again.

2.4.3 Discussion on the Assumptions of this Section

Assumptions 2 and 3. As previously discussed in Section 2.3.1, Assumption 2 aligns our stochastic game from Section 2.2 with a key feature of the dynamic Bertrand competition model: the existence of both a Nash equilibrium price and a collusive-enabling price. This assumption is also satisfied by other models, such as those of platform competition in two-sided markets (Chica et al., 2023a). Assumption 3-(i) turns our stochastic game into an infinite repeated game, where the same one-stage game is played at every stage, although firms are allowed to use one-memory strategies that condition on past price choices.

Assumption 4. This assumption is somewhat harder to interpret: part (i) is standard in the Q-learning literature, while part (ii) is used in the proof of Theorem 4 to ensure convergence of the Q-learning algorithm with bounded memory. The following sequence satisfies Assumption 4 (see Appendix Appendix A.9): Let $\alpha_1 \in [0, 1)$ be any real number and for each $k \geq 2$,

$$\alpha_k = \frac{\delta_i \alpha_{k-1}}{1 + \delta_i(1 - \delta_i)\alpha_{k-1}}.$$

Then, the sequence $\{\alpha_k\}_{k=1}^{\infty}$ satisfies Assumption 4. Moreover,

$$\alpha(\delta_i) = \frac{1}{1 - \delta_i}. \tag{2.31}$$

When (2.31) is combined with (2.28), we obtain that $Q^{i*}_{\epsilon \to 0}(\boldsymbol{p}^C, p^C) = (1 - \delta_i)^{-1}\pi^i(\boldsymbol{p}^C)$, which coincides with the value of the stochastic game when all firms play with the grim trigger strategy (see (2.21)).

Algorithm 4. In Q-learning with bounded experimentation, firms use the Q-learning algorithm with softmax exploration up to time T, which is one of the most common versions of the algorithm. After time T, firms stop exploring via softmax and begin following the argmax rule defined by the Q-function, with no further experimentation. In practice, this is the version typically used, since it is not feasible to run the softmax-based algorithm indefinitely.

Assumptions (i) and (ii) in Theorem 4. Assumption (i) Theorem 4 says that for two states, the previous price choice $\boldsymbol{p}_{T-1}$ and $\boldsymbol{p}^C$, the Q-function already weighs more the collusive-enabling price than any other price. Assumption (ii) in Theorem 4 imposes an upper bound on the value of the Q-function at time T when the previous price choice is $\boldsymbol{p}^C$. This upper bound is $(1 - \delta_i)^{-1}\pi^i(\boldsymbol{p}^C)$, which is the value of the stochastic game when all firms play with the grim trigger strategy (see (2.21)).

2.5 Conclusion

We provided a setting of stochastic games with bounded memory where one-memory SPEs exist, and formulated a version of Q-learning with bounded experimentation. We highlight our key findings:

1. We extend the theory of Fink (1964) to the case of stochastic games with bounded memory and show the existence of one-memory SPEs. We also formulate an algorithm to check whether a given profile is a one-memory SPE.

2. We formulate a version of Q-learning with bounded experimentation and show that, in stochastic games where a one-stage Nash equilibrium price and a collusive-enabling price exist, and the Q-function satisfies certain inequalities at time T, Q-learning firms charge supracompetitive prices in the long run.

3. We provide sufficient conditions under which these supracompetitive prices are supported by: (i) naive collusion, where firms always choose the collusion-enabling price; (ii) grim trigger strategies, where Q-learning firms learn to reward and punish; or (iii) increasing strategies, where firms gradually converge to the collusive-enabling price while using the Nash equilibrium price as a threat.

To our knowledge, this is the first theoretical result showing how collusion can be sustained by Q-learning firms in repeated Bertrand competition. Future work may extend our results to the case of Q-learning without bounded experimentation. We believe our setting of stochastic games with bounded memory may still be relevant to explore this extension.

Chapter 3

Competition and Collusion in Two-Sided Markets with an Outside Option

3.1 Introduction

Platform businesses have immensely grown in the last several decades due to the adoption of communication technologies.[1] For example, the sales of Amazon, which is a platform, have grown from $148 millions in 1997 to $386 billions in 2020 (Wells et al., 2021). Platforms facilitate the interaction between different types of users, such as buyers and sellers (Amazon and eBay), drivers and riders (Uber and Lyft) and content creators and consumers (YouTube, Twitch and Spotify). Their business model has become very popular, but its careful study is still in an early stage, where the first research works are from the beginning of this century (see, e.g., Rochet and Tirole (2003) and Caillaud and Jullien (2003)). There are still many open questions and, in particular, a complete model of platform competition is still far from reach.

A common yet limiting assumption in platform modeling is full market coverage, meaning that in equilibrium all users join at least one platform. While this assumption is strong, most models incorporate it because it leads to explicit equilibrium pricing formulas (see,

[1] According to the United States Census Bureau, the percentage of US citizens reporting owning a computer has grown from 8% in 1984 to 89% in 2016 (see, e.g., Ryan (2017)).

e.g., Tan and Zhou (2021)). In this work, we relax this assumption by considering a model in which N horizontally differentiated platforms compete across two market sides—buyers and sellers (collectively referred to as users)—who can either join one of the platforms (single-homing) or choose not to participate, a choice referred to as the outside option.

Dating apps offer a clear example of a market with a significant outside option and platform competition. Many users still prefer traditional non-priced methods of meeting partners (such as meeting at a bar), underscoring the importance of the outside option. This market also includes numerous competing platforms. While users typically multi-home across apps (see, e.g., the last table in SSRS (2024)[2]), some platforms encourage behavior closer to single-homing.[3] For tractability, some economic models assume single-homing in this context (see, e.g., Halaburda et al. (2018); Gal-Or (2020)). We thus use this market to ground some of our theoretical results. Ride-sharing services (e.g., Uber and Lyft) offer another example of a market with a strong outside option—public transportation, scooters, or e-bikes—but with pronounced multi-homing, as users switch between platforms depending on availability, pricing, or convenience.

We develop pricing formulas for our model and express them in terms of the equilibrium normalized net deterministic utility that platforms provide to users, i.e., the difference between the deterministic utility of users joining one platform and the deterministic utility of the outside option. This allows us to transition from a space of prices to a space of utilities in the spirit of Armstrong and Vickers (2001).

We utilize these pricing formulas to study competition and collusion between platforms with an outside option. We first establish sufficient conditions for the existence and uniqueness of a symmetric Nash equilibrium and a collusive equilibrium in this setting. We

[2]The table shows the percentage of people in different age groups who have ever used specific dating sites or apps—such as Tinder—among those who have used any dating site or app. It indicates that, on average, individuals have used about two different apps, with this average decreasing with age. However, since people may use different apps at different times, the number of apps used simultaneously is likely lower.

[3]For example, Hinge markets itself as "designed to be deleted," while loyalty-encouraging subscription models, like Bumble Boost, and curated apps, like The League or JDate, may encourage users to stick with a single app.

further show that under small cross-side externalities,[4] the normalized net deterministic utilities and market participation are smaller in collusion than in competition; furthermore, the prices on both sides are bigger in collusion than in competition. We also study the effect of increasing competition on the outputs of the competitive Nash equilibrium. In this case of increasing competition, we specify regimes for the decrease or increase of both prices and consumer surplus (these regimes depend on the size of the user's heterogeneity in tastes, the number of platforms and the size of network externalities). We demonstrate how these results can shed light on the pricing strategies observed in dating apps and how they change depending on the heterogeneity of the population and the size of network externalities.[5] We further show in this case that market participation always increases, and profits decrease if the net normalized deterministic utility is sufficiently small and increase if the net normalized deterministic utility is sufficiently large.

The size of the outside option determines the sign of the net deterministic utility in a nonlinear fashion. Indeed, we show that there exists a critical threshold for the deterministic outside option utility such that above this threshold users only receive negative deterministic utility, and below this threshold the sign of the net deterministic utility depends on the relative size of the heterogeneity in user's tastes versus the within-side externalities.

Moreover, we show that when the outside option increases, prices and consumer surplus may increase or decrease, based on the relative size of the heterogeneity in user's tastes versus the within-side externalities. In particular, we show that a model of platform competition that omits the outside option may overestimate or underestimate the true equilibrium price.

Our pricing formulas imply the following standard results for platforms, accounting

[4]Cross-side externalities capture the benefits that users on one side of the market derive from interaction with users on the other sides of the market. When these externalities are positive, platforms are confronted with the "chicken & egg" problem: to attract buyers, the platform must have a large base of sellers, who will join the platform if and only if there are many buyers in the platform (see Caillaud and Jullien (2003)).

[5]Dating markets with heterogeneous populations are dominated by general apps like Tinder and Bumble, while more homogeneous populations tend to use niche apps like The League and JDate.

for an outside option (see further references and details in Section 3.3): (i) platforms hold market power and charge users in a way that is directly proportional to user's heterogeneity in tastes; (ii) if the within-side externalities[6] are positive, platforms subsidize users on one side of the market by an amount that is proportional to the joining population on this side of the market; (iii) if the cross-side externalities are positive, platforms subsidize users on one side of the market with an amount proportional to the joining population on the other side of the market. The alignment of our results with existing ones suggests that the standard platform pricing strategy can be more general than previously understood.

Related Literature. The study of two-sided markets has emerged in the last few decades. The earlier works of Rochet and Tirole (2003, 2006), Caillaud and Jullien (2003) and Armstrong (2006) laid out the modeling foundations of two-sided markets with network externalities. These works shed some light on how equilibrium outputs of platform competition and platform monopoly depend on: (i) the size of the network externalities relative to user's heterogeneity in tastes; (ii) users being able to join either one or two platforms (i.e., having a single-home or multi-home option). Weyl (2010) and White and Weyl (2016) placed an emphasis on platform competition with insulated tariffs, which allow platforms to choose participation rates rather than prices.

Tan and Zhou (2021) modeled competition between $N \geq 2$ platforms serving multiple sides of a market. In this setting, customers are heterogeneous and modeled through random idiosyncratic preferences. Under general conditions for the probability distribution of idiosyncratic preferences, they characterized a symmetric subgame perfect equilibrium. Our model generalizes their approach by incorporating an outside option and conducting an extensive analysis for specific probability distributions. This generalization enables us to study the effects of the outside option on equilibrium outputs and collusion in two-sided

[6]Within-side externalities capture negative within-side competition effects and positive collaboration effects between users on the same side of the market. For example, competition between content creators in the same platform (e.g., *YouTube*, *Twitch* and others) and collaboration between open source programmers (e.g., *C++*, *Python*, *Linux* and others).

markets. We remark that Tan and Zhou (2021) pricing formulas are similar to the ones in Armstrong (2006), White and Weyl (2016), Jullien and Pavan (2019) and Chica et al. (2021) in the sense that they explicitly depend on the parameters. Due to the challenges of the outside option in our model, our pricing formulas are implicitly determined by the equilibrium net deterministic utility and cannot be explicit like the previous formulas. Similar to Tan and Zhou (2021), but under the assumption of an outside option, we find that increasing competition can either raise or lower equilibrium prices and consumer surplus—depending on the relative strength of network externalities versus user heterogeneity in tastes—and may also increase or decrease platform profits, depending on a condition involving the normalized net deterministic utility. However, we show that market participation always increases, whereas Tan and Zhou (2021) assumes 100% market participation. Moreover, our collusion analysis is novel and cannot be addressed in markets with full participation.

Cohen and Zhang (2022) focused on the particular setting and idiosyncrasies of ride-sharing services (e.g., *Uber* and *Lyft*). One of their interesting results is that under collusion riders pay a larger price and drivers receive a lower wage than under competition. This result is similar to one of our results (to see this one should note that the wage in their model is a negative price in our model). Nevertheless, the modeling choice for customer demand is different in both works. In their model, network externalities are endogenous, whereas ours are exogenous and thus follow the traditional setting that stems from Armstrong (2006). As such, we can examine the effects of the network externalities on the equilibrium outputs. Another major difference between these works is the type of solution obtained. They inductively solve a sequence of problems, which approximates the Nash equilibrium, and their final solution is a limit of the former solutions. We characterize the best-response of a platform that deviates from the symmetric Nash equilibrium and directly study properties of this equilibrium. The major advantage of our approach is that it allows us to study the effects of competition on the equilibrium outputs, because we can differentiate these outputs.

Our work also pertains to the literature on platform collusion. On the theoretical side, Dewenter et al. (2011) compared between competition, semi-collusion (i.e., collusion in only one side of the market) and full collusion in a special model for the newspapers market, where the two sides of the market were represented by advertisers and readers. Comparing full collusion to competition, they found out that for the advertisers market, participation is lower and prices are higher under full collusion. This result is similar to one of ours, as we show that for both sides of the markets, all sizes of within-side externalities and relatively small cross-side externalities, collusion always leads to less market participation and higher prices.[7]

Other relevant works that include an outside option are: Jeitschko and Tremblay (2020), which studied how consumers and firms endogenously choose between different homing options or the outside option; Correia-da Silva et al. (2019), which examined the welfare effects of horizontal mergers between multi-sided platforms while incorporating an outside option for consumers; Tremblay et al. (2023), which analyzed Cournot platform competition in two-sided markets with indirect network effects, where both consumers and sellers have an outside option; Peitz and Sato (2023) studied a model of asymmetric platform oligopoly while allowing partial user participation, i.e., outside options; and Teh et al. (2023) study the effects of allowing multi-homing for both sides of the market while also incorporating outside options.

Organization of the Article: Section 3.2 formulates our platform models of competition and collusion. Section 3.3 solves the models via backward induction. Section 3.4 compares the outputs of colluding and competing market models. Section 3.5 examines the effects of increasing competition on the equilibrium quantities of the competition model.

[7]Note that we focus on the study of full collusion, whereas Dewenter et al. (2011) and Lefouili and Pinho (2020) also study semi-collusion in two-sided markets and show that if the cross-side externalities are positive and sufficiently large, then semi-collusion may benefit users on the collusive side and harm users on the competitive side. Furthermore, Dewenter et al. (2011) also show that on the readers side, the collusion price may be lower than the competitive price if the competition in the advertising market is high and the newspaper market is large. We remark that our demand specification is different and, in particular, we do not incorporate a parameter for the relative size of the markets.

Section 3.6 examines the economic implications of our main results. Section 3.7 concludes this work.

3.2 The Platform Model with an Outside Option

We formulate a platform competition model by following previous works, such as White and Weyl (2016) and Tan and Zhou (2021). Let $N \geq 2$ be the number of horizontally differentiated platforms in the market. Each index $i \in \mathcal{N} := \{1, \ldots, N\}$ represents a platform competing in two different sides of a market. We index each side by $k \in \{b, s\}$, where b and s represent buyers and sellers, respectively. Each platform i sets prices for each side of the market, denoted by $\boldsymbol{p}^i = (p_b^i, p_s^i)$. The endogenous mass of users on each side of the market subscribed to platform i is denoted by $\boldsymbol{x}^i = (x_b^i, x_s^i) \in [0,1]^2$. For $i = 0$, $\boldsymbol{x}^0 = (x_b^0, x_s^0) \in [0,1]^2$ denotes the mass of users not participating in the market.

Users on side k of the market have idiosyncratic preferences for platforms and for the outside option. These preferences are captured by the i.i.d. random variables $\varepsilon_k^i \sim F_k(\cdot)$, where $k \in \{b, s\}$, $i \in \mathcal{N} \cup \{0\}$ ($i \in \mathcal{N}$ for the platforms and $i = 0$ for the outside option) and $F_k(\cdot)$ is a differentiable probability distribution.

The game consists of two stages. In stage 1, platforms strategically choose prices to maximize profits. In this article, we study two scenarios in stage 1: (i) The competition scenario, where firms compete and maximize individual profits; (ii) The collusion scenario, where firms collude and jointly maximize profits. In stage 2, given the prices determined by the platforms, users on each side of the market choose whether to participate or not and if they participate they also choose which platform to join. The game is solved using backward induction and we thus first describe the details of the second stage and then the first one.

(i) **Stage 2 (users' interactions):** Any user on side $k \in \{b, s\}$ of the market who joined platform $i \in \mathcal{N}$ receives the following utility:

$$\hat{u}_k^i := \varepsilon_k^i - p_k^i + \phi_k(\boldsymbol{x}^i), \tag{3.1}$$

where ε_k^i represents the idiosyncratic utility the user enjoys; p_k^i is the price paid by the user to access services provided by the platform (it was determined in stage 1); and $\phi_k : [0,1]^2 \longrightarrow \mathbb{R}$ is a Lipschitz differentiable function so that $\phi_k(\boldsymbol{x}^i)$ captures the network benefits that the user receives from all other users who are also joining platform i. We further denote the deterministic component of the utility by

$$u_k^i := -p_k^i + \phi_k(\boldsymbol{x}^i).$$

If a user does not join any platform, it receives the utility

$$\hat{u}_k^0 = \varepsilon_k^0 + u_k^0, \tag{3.2}$$

where $u_k^0 \in \mathbb{R}$ is a constant representing the deterministic outside utility.[8] Note that users will choose the platform that maximizes their utility, i.e., they will join platform $j \in \text{argmax}_{i \in \mathcal{N} \cup \{0\}} \{\hat{u}_k^i\}$. It follows that the mass of users from side k joining platform i solves the equation

$$x_k^i = \mathbb{P}\left(\hat{u}_k^i > \max_{j \in \mathcal{N} \cup \{0\} \setminus \{i\}} \{\hat{u}_k^j\}\right) \quad \forall k \in \{b, s\},\ i \in \mathcal{N} \cup \{0\}. \tag{3.3}$$

Proposition 8 below implies that (3.3) has at least one solution for any set $\{(p_b^i, p_s^i)\}_{i=1}^N$, and also establishes a sufficient condition for a unique solution of (3.3).

(ii) **Stage 1 (platforms' optimization):** We consider the two scenarios of competing

[8] Most models of platform competition leave out the analysis of the outside utility option that users have. By doing so, they cut out from the profit maximization process the trade-off between market participation and competition. In this article, we study this trade-off.

and colluding markets, which we describe next.

(a) A competing market: Platform i, for each $i \in \mathcal{N}$, sets the prices $\{p_b^i, p_s^i\}$ that maximize individual profits, i.e., platform i solves

$$\max_{\{p_b^i, p_s^i\}} \pi^i \left(p_b^i, p_s^i\right), \text{ where } \pi^i \left(p_b^i, p_s^i\right) := x_b^i p_b^i + x_s^i p_s^i, \tag{3.4}$$

and x_k^i is implicitly defined by (3.3). We remark that we made the common implicit assumption that the marginal costs of serving users on side b and s are zero. A Nash equilibrium associated with (3.4) is referred to as a *competitive Nash Equilibrium* (CNE).

(b) A colluding market: The colluding platforms act as a single platform trying to maximize joint profits across all sides of the market by charging one price on every side of the market; i.e., they solve

$$\max_{p_b, p_s} \Pi_{\text{tot}}(p_b, p_s), \text{ where } \Pi_{\text{tot}}(p_b, p_s) := \sum_{i=1}^{N} \left(x_b^i p_b + x_s^i p_s\right). \tag{3.5}$$

As in the competitive case, we assumed that the marginal costs of serving sellers and buyers are zero. We refer to any maximizer of (3.5) as *collusive equilibrium* (CE).[9]

In order to fully quantify the equilibrium outcomes, we make two additional assumptions:

(a) Assumption I: The idiosyncratic preferences, $\{\varepsilon_k^i\}_{k \in \{b,s\}, i \in \mathcal{N} \cup \{0\}}$, are i.i.d. Gumbel distributed with parameters (μ_k, β_k). That is, for $k \in \{b, s\}$, $\beta_k > 0$ and $\mu_k \in \mathbb{R}$, the distribution $F_k(\cdot)$ is

$$F_k(z) = e^{-e^{\frac{\mu_k - z}{\beta_k}}}. \tag{3.6}$$

We claim that this assumption is natural since it gives rise to the classical Logit model

[9] While, in general, collusion can be any situation where two or more platforms jointly make decisions, in this article, we focus on the worst-case-scenario, where all platforms collude.

(see Werden et al. (1996), Anderson and De Palma (1992)), which describes the demand of heterogeneous consumers for a set of differentiated goods (see Berry (1994), Conlon and Gortmaker (2020), Besanko et al. (1998)). To support this claim, we note that the central equation in this work is (3.3), which can be rewritten using (3.2) and the alternative variables $\theta_k^i = \varepsilon_k^i - \varepsilon_k^0$, $i \in \mathcal{N} \cup \{0\}$, $k \in \{b, s\}$ as follows:

$$x_k^i = \mathbb{P}\left(\theta_k^i + u_k^i \geq \max_{j=0,1,\ldots,N, j\neq i} \left\{\theta_k^j + u_k^j\right\}\right), \quad \forall k \in \{b, s\},\ i \in \mathcal{N} \cup \{0\}.$$

We further note that $\theta_k^i \sim \text{Logistic}(0, \beta_k)$ for $i \in \mathcal{N}$ and thus conclude the claim.

(b) Assumption II: The function $\phi_k(\boldsymbol{x})$ is linear, i.e., it can be represented by multiplying with the real-valued matrix $\Phi \in \mathbb{R}^{2\times 2}$

$$\begin{bmatrix} \phi_b(x_b, x_s) \\ \phi_s(x_b, x_s) \end{bmatrix} = \underbrace{\begin{bmatrix} \phi_{bb} & \phi_{bs} \\ \phi_{sb} & \phi_{ss} \end{bmatrix}}_{\Phi} \begin{bmatrix} x_b \\ x_s \end{bmatrix}. \tag{3.7}$$

We remark that the following results do not require these additional assumptions: The existence and uniqueness of the solution of (3.3) (see Proposition 8 in Section 3.3) and the derivation of the first- and second-order conditions for both (3.4) and (3.5) (see Lemma 2, 3, 4 and 5 in Appendix Appendix B). We note that many of the results presented in this paper can be extended to other probability distributions of economic interest. In the Online Appendix, we demonstrate this for the exponential distribution with two platforms. In particular, we explicitly derive the first-order condition for (3.4) using Mathematica. Additionally, we present numerical simulations supporting results similar to those shown in Propositions 14, 15, and 16, but using the exponential distribution. It is important to recognize that each probability distribution requires special treatment, and the analysis of the Gumbel distribution is already quite lengthy and complex.

3.3 Equilibrium

We solve our model using backward induction. We first study the solution to (3.3) and show that for any set of prices, $\{(p_b^i, p_s^i)\}_{i=1}^N$, there is a well-defined set of market shares, $\{(x_b^i, x_s^i)\}_{i=0}^N$, that solve (3.3) and under a certain condition they are unique. Next, we characterize the symmetric CNE of (3.4) (i.e., the CNE such that $p_k^i = p_k^*$ for each $i \in \mathcal{N}$) and the CE of (3.5). At last, we interpret the resulting equilibrium pricing and market share formulas.

Stage 2 Solution: Users' Maximization

We establish sufficient conditions for the existence and uniqueness for (3.3). We recall that (3.3) captures the users' dynamics when prices change. Let $\boldsymbol{u} = (u^0, u^1, \ldots, u^N) \in \mathbb{R}^{N+1}$ and for $k \in \{b, s\}$ and $i \in \mathcal{N} \cup \{0\}$ define

$$T_k^i(\boldsymbol{u}) := \mathbb{P}(\epsilon_k^i > \max_{j \neq i}(\epsilon_k^j + u^j - u^i)). \tag{3.8}$$

In view of (3.8), (3.3) can be rewritten as

$$x_k^i = T_k^i(u_k^0, \phi_k(\boldsymbol{x}^1) - p_k^1, \ldots, \phi_k(\boldsymbol{x}^N) - p_k^N). \tag{3.9}$$

It follows that a vector $\boldsymbol{x} = (x_b^0, x_s^0, \ldots, x_b^N, x_s^N)$ solves (3.3) if and only if it is a fixed point of the map $\Sigma : [0,1]^{2(N+1)} \longrightarrow [0,1]^{2(N+1)}$ given by

$$\Sigma(\boldsymbol{x}) := (T_b^0(\boldsymbol{u}_b), T_s^0(\boldsymbol{u}_s), \ldots, T_b^N(\boldsymbol{u}_b), T_s^N(\boldsymbol{u}_s)), \tag{3.10}$$

where $\boldsymbol{u}_k = (u_k^0, \phi_k(\boldsymbol{x}^1) - p_k^1, \ldots, \phi_k(\boldsymbol{x}^N) - p_k^N)$ and $\boldsymbol{x}^i = (x_b^i, x_s^i)$. Proposition 8 below, shows that (3.10) always has at least one fixed point and thus (3.3) always has a solution. This proposition also provides sufficient conditions for the uniqueness of this fixed point and the solution of (3.3). Its formulation requires the following Lipschitz-type constants:

$$\begin{aligned} M_T &:= \max_{k\in\{b,s\}, i\in\mathcal{N}\cup\{0\}} \sup_{\boldsymbol{u}\in\mathbb{R}^{N+1}} \sum_{l=0}^{N} \left| \frac{\partial T_k^i(\boldsymbol{u})}{\partial u^l} \right|, \text{ and} \\ M_\phi &:= \max_{k\in\{b,s\}} \sup_{(x_b,x_s)\in[0,1]^2} \sum_{l\in\{b,s\}} \left| \frac{\partial \phi_k(x_b,x_s)}{\partial x_l} \right|. \end{aligned} \tag{3.11}$$

We remark that $\partial T_k^i(\boldsymbol{u})/\partial u^l$ captures the user's sensitivity to changes in utility levels. Similarly, $\partial\phi_k(x_b,x_s)/\partial x_l$ measures how externalities change when more people join one specific platform.

Proposition 8 (Existence and Uniqueness of Market Shares)**.** *For any prices $\{(p_b^i, p_s^i)\}_{i=1}^N \subset \mathbb{R}^2$, there exists a solution to (3.3), $\boldsymbol{x} = (x_b^0, x_s^0, x_b^1, x_s^1, \cdots, x_b^N, x_s^N)$, such that for each $k \in \{b,s\}$, $\sum_{i=0}^N x_k^i = 1$. Moreover, if $M_T M_\phi < 1$, where M_T and M_ϕ are given by (3.11), then the solution of (3.3) is unique.*

This Proposition provides sufficient conditions for the mapping from prices to market shares, $\{(p_b^i, p_s^i)\}_{i=1}^N \mapsto \{(x_b^i, x_s^i)\}_{i=0}^N$, to be well-defined. Its proof is in Appendix Appendix B.[10]

Stage 1 Solution: Platforms' Optimization

We establish sufficient conditions for the existence and uniqueness of symmetric solutions of (3.4) and (3.5). We first focus on symmetric solutions for (3.4). For this purpose, we use the following transformation:

$$z_k := \frac{u_k - u_k^0}{\beta_k}, \quad \text{for } k \in \{b,s\}. \tag{3.12}$$

[10] We remark that while we use below Proposition 8 to find symmetric Nash equilibria, this proposition cannot be restricted for symmetric market shares. When proving the existence of symmetric Nash equilibria (see Proposition 10), one needs to consider all possible deviations from the equilibrium path, in particular, those that are not necessarily on the symmetric path.

We note that $u_k - u_k^0 \equiv -p_k + \phi_k(\boldsymbol{x}) - u_k^0$ captures the difference between the deterministic utility of users (sellers or buyers) joining one platform and the deterministic utility of the outside option. We remark that in the symmetric case any platform charges the price p_k for $k \in \{b, s\}$ and the market shares are given by $\boldsymbol{x} = (x_b, x_s)$. The Gumbel distribution parameter β_k is a measure of the standard deviation of the idiosyncratic preference ε_k^i and it captures the degree of heterogeneity in users' tastes.[11] Throughout the article, we will refer to z_k as the *normalized net deterministic utility* of users on side k of the market.

We can write the first-order condition (FOC) of (3.4) as a function of z_k.[12]

Proposition 9 (FOC of (3.4))**.** *Suppose there is a symmetric equilibrium (p_b^*, p_s^*) solution of (3.4) with market shares (x_b^*, x_s^*). If one platform unilaterally deviates from this symmetric CNE, the FOC that characterizes its best-response is given by*

$$\boldsymbol{\beta z} = (\Phi - H(\boldsymbol{z}))\,\Omega(\boldsymbol{z}) - \boldsymbol{u}_0, \tag{3.13}$$

where $\boldsymbol{\beta} = \mathrm{diag}(\beta_b, \beta_s)$, $\boldsymbol{z} = (z_b, z_s)$, $\boldsymbol{u}_0 = (u_b^0, u_s^0)$, Φ is the externalities network matrix defined in (3.7), $\Omega(\boldsymbol{z}) = (\omega(z_b), \omega(z_s))^T$ with $\omega : \mathbb{R} \longrightarrow (0, \frac{1}{N})$ such that $\omega(z) := \frac{1}{e^{-z}+N}$, and $H(\boldsymbol{z})$ is a $2x2$ matrix defined as

$$H(\boldsymbol{z}) := \begin{bmatrix} L_b d_b K_s + h_b - \phi_{bb} & -\phi_{sb}(d_s L_b + 1) \\ -\phi_{bs}(d_b L_s + 1) & L_s d_s K_b + h_s - \phi_{ss} \end{bmatrix}, \tag{3.14}$$

[11] Note that in general, if $\varepsilon_k^i \sim G(\mu_k, \beta_k)$, then $\mathrm{Var}[\varepsilon_k^i] = \frac{\pi^2}{6}\beta_k^2$ and thus the standard deviation of ε_k^i is $\frac{\pi}{\sqrt{6}}\beta_k$.

[12] It is a known fact that attempting to solve (3.4) by means of an FOC with respect to prices $\{(p_b^i, p_s^i)\}_{i=1}^N$ produces a non-tractable system of equations (see, e.g., Tan and Zhou (2021) and Chica et al. (2021)). By contrast, our proofs in the appendix take derivatives with respect to $\{x_b^i, x_s^i\}_{i=1}^N$. By Proposition 8 and the implicit function theorem, there is a well-defined locally 1-1 mapping from $\{(x_b^i, x_s^i)\}_{i=1}^N$ to $\{(p_b^i, p_s^i)\}_{i=1}^N$.

where L_k, d_k and h_k for each $k \in \{b, s\}$ are given by

$$\begin{aligned} L_k &= \frac{(N-1)\beta_k}{J_\phi}(1 + Ne^{z_k}), \\ d_k &= \beta_k(1 + Ne^{z_k}), \\ h_k &= \beta_k(1 + e^{z_k})(e^{-z_k} + N), \\ K_k &= \phi_{kk} - \beta_k(1 + Ne^{z_k})(e^{-z_k} + N - 1), \\ J_\phi &= K_b K_s - \phi_{sb}\phi_{bs}. \end{aligned} \tag{3.15}$$

Let us assume $\boldsymbol{z}^* = (z_b^*, z_s^*)^T$ is the unique solution of (3.13) and it satisfies a corresponding second order condition. We discuss below (see Proposition 10) sufficient conditions for this assumption. We use $\boldsymbol{z}^*$ to characterize the symmetric equilibrium solution of (3.4), $\boldsymbol{p}^* = (p_b^*, p_s^*)^T$, with market shares $\boldsymbol{x}^* = (x_b^*, x_s^*)^T$. By applying (3.8) and (3.9) evaluated at $u_k^i = u_k^* = -p_k^* + \phi_k(\boldsymbol{x}^*)$, where $i \in \mathcal{N}$ and $k \in \{b, s\}$, one can show (see (B.40) in Appendix Appendix B) that

$$x_k^* = \omega(z_k^*) \equiv \frac{1}{e^{-z_k^*} + N} \quad \text{and thus} \quad \boldsymbol{x}^* = \Omega(\boldsymbol{z}^*). \tag{3.16}$$

We further note that (3.12) implies that $\boldsymbol{\beta}\boldsymbol{z}^* = -\boldsymbol{p}^* + \Phi\boldsymbol{x}^* - \boldsymbol{u}_0$. Combining the latter equation, (3.13) and (3.16), the symmetric CNE of (3.4) is given by

$$\begin{aligned} \boldsymbol{p}^* &= H(\boldsymbol{z}^*)\Omega(\boldsymbol{z}^*) \text{ and} \\ \boldsymbol{x}^* &= \Omega(\boldsymbol{z}^*). \end{aligned} \tag{3.17}$$

In order to ensure that (3.17) yields the symmetric CNE, we next establish a sufficient condition for (3.13) to have a unique solution that satisfies a corresponding second order condition. It uses the following function

$$f(N) := \frac{2(N-1)}{N^2}, \tag{3.18}$$

where we note that f approaches 0 as $N \to \infty$. It also uses the notation $B_\epsilon(0)$ for the ball in $\mathbb{R}^2$ of radius $\epsilon > 0$ around the origin.

Proposition 10 (Existence and uniqueness of the symmetric CNE)**.** *Suppose that $N \geq 2$ and for each $k \in \{b, s\}$, (ϕ_{kk}, β_k) satisfies*

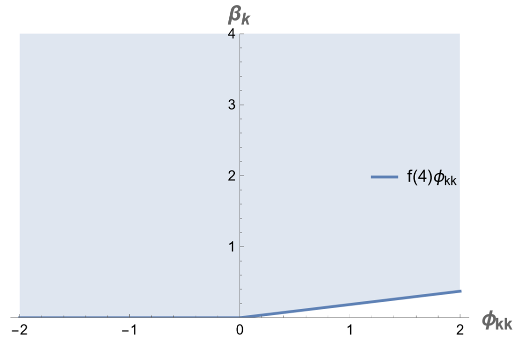
$$\text{either } (\phi_{kk} \leq 0 \text{ and } \beta_k > 0) \text{ or } (\phi_{kk} > 0 \text{ and } \beta_k > f(N)\, \phi_{kk}). \quad (3.19)$$

Then, there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$ there is a unique solution of (3.13) and this solution satisfies a second order condition.[13] *Furthermore, (3.17) yields the unique symmetric CNE of (3.4).*

Proposition 10 guarantees the existence and uniqueness of a symmetric CNE for a large family of the parameters $\{\phi_{kk}, \beta_k\}_{k\in\{b,s\}}$. In particular, if the within-side externalities (i.e., those that reflect interactions of the same sides of the market), ϕ_{kk}, are negative, then existence and uniqueness of a solution for (3.13) is guaranteed for any size of heterogeneity in user's tastes, β_k. On the other hand, if the within-side externalities are positive, then existence and uniqueness is only ensured for relatively large sizes of heterogeneity in user's tastes (i.e., $\beta_k > f(N)\phi_{kk}$). Recall that as the number of platforms N grows to infinity, $f(N)$ approaches 0. Thus, even for positive within-side externalities, existence and uniqueness of a solution for (3.13) is guaranteed for any size of β_k, provided that the number of platforms in the market is large enough. Some form of the latter condition appears in many studies of platform competition (see, e.g., Anderson et al. (1992), Armstrong (2006), and Tan and Zhou (2021)). This condition ensures that network effects do not always dominate idiosyncratic preferences when users are charged non-zero prices (see, e.g., Chica et al. (2021)). Figure 3.1 below shows the region described by (3.19) when $N = 4$.

[13] We clarify that the ϵ in Proposition 10 depends on $(\phi_{bb}, \phi_{ss}, \beta_b, \beta_s, N, u_b^0, u_s^0)$, but for simplicity we denote it by ϵ. We use the same convention in other places in this article where a similar condition with ϵ appears.

Figure 3.1: Region of (ϕ_{kk}, β_k) that guarantees a unique symmetric CNE when $N = 4$ according to Proposition 10.



Next, we focus on the solution of (3.5). We first establish the FOC of (3.5) as a function of z_k.

Proposition 11 (FOC of (3.5))**.** *The FOC of* (3.5) *is given by*

$$\boldsymbol{\beta}\boldsymbol{z} = (\Phi - H^C(\boldsymbol{z}))\Omega(\boldsymbol{z}) - \boldsymbol{u}_0, \tag{3.20}$$

where $\boldsymbol{\beta}$, $\boldsymbol{z}$, $\boldsymbol{u}_0$, Φ, $\Omega(\boldsymbol{z})$ were defined in Proposition 9, and $H^C(\boldsymbol{z})$ is a $2x2$ matrix defined by

$$H^C(\boldsymbol{z}) := \begin{bmatrix} \frac{\beta_b(1+Ne^{z_b})^2}{e^{z_b}} - \phi_{bb} & -\phi_{sb} \\ -\phi_{bs} & \frac{\beta_s(1+Ne^{z_s})^2}{e^{z_s}} - \phi_{ss} \end{bmatrix}. \tag{3.21}$$

Let us assume $\boldsymbol{z}^{\mathrm{C}} = (z_b^{\mathrm{C}}, z_s^{\mathrm{C}})^T$ is the unique solution of (3.20) and it satisfies a corresponding second order condition (we provide sufficient conditions for these assumptions in Proposition 12 below). Following the same derivation of (3.17) (see the proof of Proposition 11 in Appendix Appendix B), one can show that the CE solution of (3.5),

$\boldsymbol{p}^{\mathrm{C}} = (p_b^{\mathrm{C}}, p_s^{\mathrm{C}})^T$ and the corresponding market shares, $\boldsymbol{x}^{\mathrm{C}} = (x_b^{\mathrm{C}}, x_s^{\mathrm{C}})^T$, satisfy

$$\begin{aligned} \boldsymbol{p}^{\mathrm{C}} &= H^{\mathrm{C}}(\boldsymbol{z}^{\mathrm{C}})\Omega(\boldsymbol{z}^{\mathrm{C}}) \text{ and} \\ \boldsymbol{x}^{\mathrm{C}} &= \Omega(\boldsymbol{z}^{\mathrm{C}}). \end{aligned} \tag{3.22}$$

In order to ensure that (3.22) yields the CE, we establish sufficient conditions for (3.20) to have a unique solution that satisfies a corresponding second order condition.

Proposition 12 (Existence and uniqueness of the CE)**.** *For any* $u_b^0, u_s^0 \in \mathbb{R}$, $\Phi \in \mathbb{R}^{2x2}$, $\beta_b, \beta_s > 0$ *and* $N \geq 2$, *there exists a solution for (3.20). Moreover, if for each* $k \in \{b, s\}$, (ϕ_{kk}, β_k) *satisfies*

$$\text{either } (\phi_{kk} \leq 0 \text{ and } \beta_k > 0) \;\text{ or }\; (\phi_{kk} > 0 \text{ and } \beta_k > \frac{8\phi_{kk}}{27N}), \tag{3.23}$$

then there exists $\varepsilon > 0$ *such that for any* $(\phi_{bs}, \phi_{sb}) \in B_\varepsilon(0) \subset \mathbb{R}^2$, *the solution for (3.20) is unique, it satisfies a corresponding second order condition and (3.22) yields the unique CE of* (3.5).

The proof of Proposition 12 implies that $f(N)\phi_{kk}$, which was used in Proposition 10, is strictly bigger than $8\phi_{kk}/(27N)$ for all $\phi_{kk} > 0$. It follows that, if (ϕ_{kk}, β_k) satisfies (3.19) for each $k \in \{b, s\}$, then it also satisfies (3.23) and consequently there exists unique solutions $\boldsymbol{z}^*$ and z^{C} to (3.13) and (3.20), respectively. Section 3.4 will compare these two solutions assuming (3.19) is satisfied.

Interpretation and Implications of the Resulting Pricing Formulas

We discuss the pricing formulas (3.17) and (3.22) for the competition and collusion models. We first relate them to common pricing competition models. Both formulas are expressed in terms of the equilibrium normalized net deterministic utility, z_k, that platforms provide to users on both sides of the market. They thus remind the formulation in Armstrong and

Vickers (2001), where multiple firms compete in a utility space, instead of a space of prices. When solving (3.4), firms internalize competition for users in terms of the utility they can provide w.r.t. (with respect to) the outside utility. The optimal vector utility, $\boldsymbol{z}^*$, provided by the competing platforms is determined so that some users are always excluded from the market.[14] A similar result is obtained for the colluding case, while excluding a larger portion of participants, as shown below in Proposition 18. Therefore, our models also imply the standard result that the output is not optimally distributed among users when there is price competition or collusion (see, e.g., Varian (1989), Armstrong (1996) and Rochet and Choné (1998)).

Our pricing formulas (3.17) and (3.22) generalize many of the standard results in the platform's literature for the case of an outside option. We emphasize some of these generalizations: (i) For CE, the term $\beta_k(1+Ne^{z_k})^2/e^{z_k}$, which appears in the diagonal of the matrix (3.21), captures the platform's market power (see Perloff and Salop (1985)). It implies that in equilibrium platforms charge users on side k of the market proportionally to the platform's differentiation parameter β_k (see Tan and Zhou (2021) and Chica et al. (2021)). (ii) For CNE and CE, assume that the within-side externalities are positive (i.e., $\phi_{kk} \geq 0$). Then, from the diagonal of (3.14) and (3.21), platforms subsidize users on side k by an amount that is proportional to the joining population on this side of the market (i.e., they subsidize users on side k with $\phi_{kk}x_k^*$ and $\phi_{kk}x_k^{\mathrm{C}}$, respectively, for the competing and colluding models). If these externalities are negative (i.e., $\phi_{kk} < 0$), the opposite result is true (see Bardey et al. (2014)). (iii) For CNE and CE, assume positive cross-side externalities, that is, $\phi_{lk} \geq 0$ for each $l, k \in \{b, s\}$, $l \neq k$. Then, the off-diagonal terms of (3.14) and (3.21) imply that platforms subsidize users on side k with an amount directly affected by the joining population on the other side of the market (i.e., platforms subsidize an amount $\phi_{lk}x_l^*$ to users on side k).

[14] Note that the equilibrium market share satisfies $x_k^* = \omega(z_k^*) < 1/N$ (see (3.17) and the definition of $\omega(\cdot)$ in Proposition 9). This condition excludes the participation of some users.

3.4 Competition and Collusion in Two-sided Markets with an Outside Option

We compare the colluding and competing market models by studying the main properties of and differences between the pricing formulas (3.17) and (3.22). We first assume competition and characterize the markets in which users receive positive and negative normalized net deterministic utility, z_k^* (see Proposition 13 and Corollary 1). We also characterize the sign of z_k^* under perfect competition (i.e., as $N \to \infty$) and show that platforms charge a price that is equal to user's heterogeneity in tastes while covering the entire market (see Corollary 2). We then study the effects of the outside option on the change of prices, profits and consumer surplus. In particular, we show that when the outside option increases: (i) prices on side k may increase or decrease (see Proposition 14); (ii) profits decrease (see Proposition 15); and (iii) consumer surplus may increase or decrease (see Proposition 16). Next, we assume collusion and characterize markets in which z_k^{C} is positive or negative (see Proposition 17 and Corollaries 3 and 4). Finally, we compare the equilibrium quantities of competition and collusion (see Proposition 18).

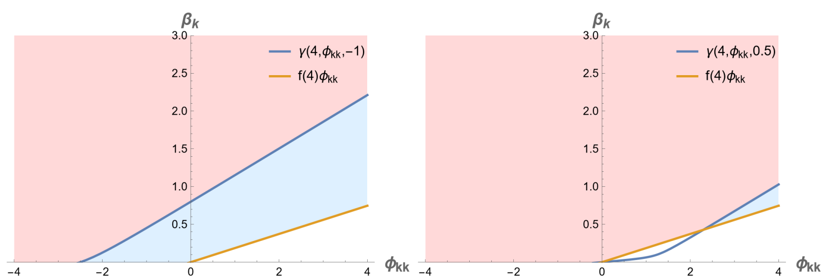
The Sign of the Net Deterministic Utility Under Competition

In CNE, a positive (negative) z_k^* implies that the deterministic utility that users enjoy in equilibrium from joining a given platform is larger (smaller) than the deterministic utility of the outside option. For this reason, we first study the sign of z_k^* as given by the solution of (3.13). The following proposition shows sufficient conditions to partition the region described by (3.19) into two regions: $\{z_k^* < 0\}$ and $\{z_k^* > 0\}$, which we demonstrate in Figure 3.2 for two different values of u_k^0. The indifference region $\{z_k^* = 0\}$ is described by a curve $\beta_k = \gamma(N, \phi_{kk}, u_k^0)$ in the plane (ϕ_{kk}, β_k), where γ is defined as follows:

$$\gamma(N, \phi_{kk}, u_k^0) := \frac{(2\phi_{kk} - Nu_k^0) + \sqrt{(2\phi_{kk} - Nu_k^0)^2 + 4\phi_{kk}\left(u_k^0 - \frac{2\phi_{kk}}{N+1}\right)}}{2(N+1)}. \tag{3.24}$$

We remark that the clustering of the sign of z_k^* according to this proposition requires a local bound on the cross-side externalities.

Figure 3.2: Classification of the sign of z_k^* based on (ϕ_{kk}, β_k), $k \in \{b, s\}$, according to Proposition 13, where $N = 4$ and $u_k^0 = -1$ (left) or $u_k^0 = 0.5$ (right). The red and blue regions correspond to negative and positive z_k^*, respectively.



Proposition 13 (The sign of z_k^*)**.** *Suppose that $N \geq 2$ and for each $k \in \{b, s\}$, (ϕ_{kk}, β_k) satisfies (3.19). Then we can further partition the domain specified in (3.19) into two regions, which cluster the sign of z_k^* as long as a local condition on the cross-side externalities hold:*

(i) If $\beta_k > \gamma(N, \phi_{kk}, u_k^0)$, then there exists $\varepsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\varepsilon(0) \subset \mathbb{R}^2$, $z_k^* < 0$.

(ii) $\beta_k < \gamma(N, \phi_{kk}, u_k^0)$, then there exists $\varepsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\varepsilon(0) \subset \mathbb{R}^2$, $z_k^* > 0$.

We first clarify the economic meaning of this proposition. If user's heterogeneity in tastes is large enough (i.e., $\beta_k > \gamma(N, \phi_{kk}, u_k^0)$), then it is a standard result that platforms extract consumer surplus by charging a price that leads to a negative normalized net deterministic utility, i.e., $z_k^* < 0$ (see Anderson and De Palma (1992), Tan and Wright (2021), Chica et al. (2021) and others).[15] On the other hand, if the user's heterogeneity in tastes, β_k, is small enough (i.e., $\beta_k < \gamma\left(N, \phi_{kk}, u_k^0\right)$), then users receive positive normalized net deterministic utility, i.e., $z_k^* > 0$.

We identify a critical threshold for the deterministic outside utility so that above this threshold, (ii) in Proposition 13 is not feasible.

Corollary 1 (The sign of z_k^* for large values of u_k^0)**.** *Case (ii) in Proposition 13 is not feasible if $u_k^0 \geq \tilde{u}_k^0(N, \phi_{kk})$, where $\tilde{u}_k^0(N, \phi_{kk})$ is the critical threshold for the deterministic outside utility and it is defined in (B.99) in the Appendix Appendix B.*

This corollary implies that if the deterministic outside utility is sufficiently large (i.e., $u_k^0 \geq \tilde{u}_k^0(N, \phi_{kk})$), the CNE leads to a negative net deterministic utility for any size of heterogeneity in user's tastes satisfying (3.19). In other words, only if the deterministic outside utility is relatively small (i.e., $u_k^0 < \tilde{u}_k^0(N, \phi_{kk})$), users with relatively weak preferences (i.e., $\beta_k < \gamma(N, \phi_{kk}, u_k^0)$) receive positive net deterministic utility.

Next, we show that in the case of *perfect competition* (i.e., the limiting case $N \to \infty$), the sign of z_k^* can be characterized by the sign of u_k^0 and the size of β_k.

Corollary 2 (CNE under perfect competition)**.** *For each $k \in \{b, s\}$, any $u_k^0 \in \mathbb{R}$, $\Phi \in \mathbb{R}^{2x2}$ and $\beta_k > 0$, under perfect competition (i.e., when $N \to \infty$),*

$$\lim_{N\to\infty} z_k^* \begin{cases} > 0, & \text{if } u_k^0 < 0 \text{ and } \ \beta_k < -u_k^0; \\ < 0, & \text{if } (u_k^0 < 0 \text{ and } \ \beta_k > -u_k^0) \text{ or } \ u_k^0 \geq 0. \end{cases} \tag{3.25}$$

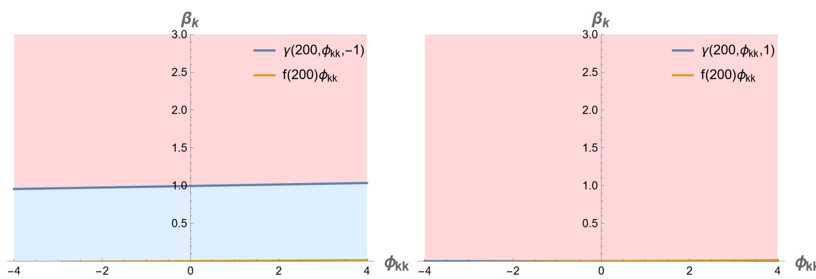
Moreover, as $N \to \infty$, $p_k^* \to \beta_k$, $x_k^* \to 0$ and $Nx_k^* \to 1$.

Under perfect competition, platforms charge a price that is equal to the user's heterogeneity in tastes for that side of the market, i.e., $p_k^* = \beta_k$. Moreover, the equilibrium market

[15] This result is due to the fact that highly heterogeneous users are less responsive to price and demand effects.

participation on side k of the market is complete, i.e., $Nx_k^* = 1$. When the deterministic outside option utility is positive, users receive negative normalized net deterministic utility for any size of β_k under perfect competition. When the deterministic outside option utility is strictly negative, users receive positive normalized net deterministic utility if and only if the heterogeneity of users' taste is small, i.e. $\beta_k < -u_k^0$. We demonstrate (3.25) in Figure 3.3 for two different values (negative and positive) of u_k^0 and a sufficiently large N.

Figure 3.3: Classification of the sign of z_k^* based on (ϕ_{kk}, β_k), $k \in \{b, s\}$, for large N according to Corollary 2, where $N = 200$ and $u_k^0 = -1$ (left) or $u_k^0 = 1$ (right). The red and blue regions correspond to negative and positive z_k^*, respectively.



The Effects of the Outside Option on the CNE

The following proposition provides sufficient conditions to characterize the sign of $\partial p_k^* / \partial u_k^0$. It shows that the effect of the outside option on p_k^* is nonlinear. It uses the following quantities:

$$
\begin{aligned}
g_{p,u}(N) &:= \frac{\left(N + \sqrt{(N-1)(N+3)} + 1\right)}{2N} \quad \text{and} \\
f_{p,u}(N) &:= \frac{1}{2}\left(\sqrt{\frac{N-2}{N}} + 1\right).
\end{aligned} \tag{3.26}
$$

Proposition 14. *(The sign of $\partial p_k^* / \partial u_k^0$) Suppose that $N \geq 2$ and for each $k \in \{b, s\}$, (ϕ_{kk}, β_k) satisfies (3.19). Then we can further partition the domain specified in (3.19) into two regions, which cluster the sign of $\partial p_k^* / \partial u_k^0$ as long as a local condition on the cross-side* *externalities hold:*

(i) If either ($\phi_{kk} \leq 0$ *and* $\beta_k > 0$) *or* ($\phi_{kk} > 0$ *and* $\beta_k > g_{p,u}(N)\,\phi_{kk}$)*, then there exists* $\varepsilon > 0$ *such that for any* $(\phi_{bs}, \phi_{sb}) \in B_\varepsilon(0) \subset \mathbb{R}^2$, $\partial p_k^* / \partial u_k^0 < 0$.

(ii) If $\phi_{kk} > 0$, $N \geq 3$ *and* $f(N)\,\phi_{kk} < \beta_k < f_{p,u}(N)\,\phi_{kk}$*, then there exists* $\varepsilon > 0$ *such that for any* $(\phi_{bs}, \phi_{sb}) \in B_\varepsilon(0) \subset \mathbb{R}^2$, $\partial p_k^* / \partial u_k^0 > 0$.

Moreover, if $(\phi_{bs}, \phi_{sb}) = 0$*, then*

$$
\begin{aligned}
&\lim_{u_k^0 \to -\infty} p_k^* = \frac{N}{N-1}\beta_k - \frac{\phi_{kk}}{N-1} =: p_{k,u}, \text{ and} \\
&\lim_{u_k^0 \to \infty} p_k^* = \beta_k =: p_{k,E}.
\end{aligned}
\tag{3.27}
$$

Consequently, $p_k^* \in (p_{k,E}, p_{k,u})$ *in case (i), and* $p_k^* \in (p_{k,u}, p_{k,E})$ *in case (ii).*

We clarify the economic meaning of this proposition. We first note it implies $p_k^* \to p_{k,u}$ when $u_k^0 \to -\infty$, which coincides with the equilibrium price in a platform competition model with no outside option. It also implies $p_k^* \to p_{k,E} = \beta_k$ when $u_k^0 \to \infty$, which represents the efficient price, that is, the price, β_k, under perfect competition, expressed in Corollary 2. In part (i), the incorporation of an outside option into the platform competition model decreases the equilibrium price w.r.t. the no outside option model, which is an expected result. Thus, if the cross-side externalities are sufficiently small and the within-side externalities are either negative or positive with relatively large user's heterogeneity in tastes, then users are compensated by an amount equal to $p_{k,u} - p_k^*$. Moreover, in this case, users always pay a price that is bigger than the efficient price, $p_{k,E}$. On the other hand, in part (ii), incorporating an outside option increases the equilibrium price w.r.t the no outside option model, which is non-trivial. Therefore, under sufficiently small cross-side externalities and positive within-side externalities, users with relatively small heterogeneity in tastes pay a premium w.r.t the model with no outside option, which is quantified by $p_k^* - p_{k,u}$. Moreover, users always pay a price that is smaller than the efficient price $p_{k,E}$.

Remark 1 (Price overestimation vs. underestimation)**.** *If a given population can be parameterized using the region of parameters described by either (i) or (ii) of Proposition 14, then a model of platform competition that omits the outside option will either overestimate or underestimate, respectively, the true equilibrium price.*

The following proposition shows sufficient conditions to determine the sign of $\partial \pi_k^* / \partial u_k^0$. It uses the following quantity:

$$g_{\pi,u}(N) := \sqrt{\frac{N-1}{N^3}} + \frac{1}{N}. \tag{3.28}$$

Proposition 15. *(The sign of $\partial \pi_k^* / \partial u_k^0$) If $N \geq 2$ and*

$$\text{either } (\phi_{kk} \leq 0 \text{ and } \beta_k > 0) \text{ or } (\phi_{kk} > 0 \text{ and } \beta_k > g_{\pi,u}(N)\, \phi_{kk}),$$

then there exists $\varepsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\varepsilon(0) \subset \mathbb{R}^2$, $\partial \pi_k^* / \partial u_k^0 < 0$. Moreover, if $(\phi_{bs}, \phi_{sb}) = 0$, then

$$\begin{aligned} \lim_{u_k^0 \to -\infty} \pi_k^* &= \frac{\beta_k}{N-1} - \frac{\phi_{kk}}{(N-1)\,N} =: \pi_{k,u}, \text{ and} \\ \lim_{u_k^0 \to \infty} \pi_k^* &= 0 =: \pi_{k,E}. \end{aligned} \tag{3.29}$$

Note that even though, by part (ii) of Proposition 14, prices may increase with the outside option utility, Proposition 15 shows that profits are always decreasing w.r.t. u_k^0. This happens because market participation is always decreasing w.r.t. u_k^0. Therefore, it is not a surprise that profits are decreasing as a function of u_k^0.

The following proposition provides sufficient conditions to determine the sign of the derivative of the consumer surplus w.r.t. the outside option utility. More specifically, it uses the equilibrium consumer surplus on side k of the market, CS_k^*, which is defined as follows (see Tan and Zhou (2021) for the case without an outside option):

$$CS_k^* := \mathbb{E}\left[\max_{i=0,\ldots,N} \epsilon_k^i\right] - p_k^* + \phi_k\left(x_b^*, x_s^*\right), \tag{3.30}$$

where $\mathbb{E}\left[\max_{i=0,\ldots,N} \epsilon_k^i\right]$ is the expected maximum idiosyncratic utility.[16] It also uses the function $f_{CS,u}(N)$ given in Appendix Appendix B (see (B.120)).

Proposition 16 (The sign of $\partial(CS_k^*)/\partial u_k^0$)**.** *The effect of the outside option on the change of consumer surplus can be clustered into the following two regions:*

(i) If $N \geq 2$ and either ($\phi_{kk} \leq 0$ and $\beta_k > 0$) or ($\phi_{kk} > 0$ and $\beta_k > 2\phi_{kk}$), then there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, $\partial(CS_k^*)/\partial u_k^0 > 0$.

(ii) If $N \geq 2$, $\phi_{kk} > 0$, $f(N)\,\phi_{kk} < \beta_k < f_{CS,u}(N)\,\phi_{kk}$, then there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, $\partial(CS_k^*)/\partial u_k^0 < 0$.

Part (i) of this proposition implies that the incorporation of an outside option into the platform competition model may increase the consumer surplus, or equivalently, the consumer welfare, w.r.t. the no outside option model. On the other hand, Part (ii) implies that incorporating an outside option may decrease the equilibrium consumer surplus w.r.t the no outside option model. While part (i) is standard, part (ii) is surprising.

The Sign of the Net Deterministic Utility Under Collusion

The following proposition quantifies the sign of z_k^{C} in the collusion case of (3.5). In particular, it claims that the indifference region $\{z_k^{\mathrm{C}} = 0\}$ is described by a curve $\beta_k = \gamma^{\mathrm{C}}(N, \phi_{kk}, u_k^0)$ in the plane (ϕ_{kk}, β_k), where γ^{C} is defined as follows:

$$\gamma^{\mathrm{C}}(N, \phi_{kk}, u_k^0) := \frac{2\phi_{kk} - u_k^0\,(N+1)}{(N+1)^2}. \tag{3.31}$$

[16]Note that $\max_{i=0,\ldots,N}\{\epsilon_k^i\} \sim G(\mu_k + \beta_k \ln(N+1), \beta_k)$ and thus $\mathbb{E}[\max_{i=0,\ldots,N} \epsilon_k^i] = [\mu_k + \beta_k \ln(N+1)] + \beta_k\gamma$, where γ denotes the Euler-Mascheroni constant. This quantity captures the product variety of the market.

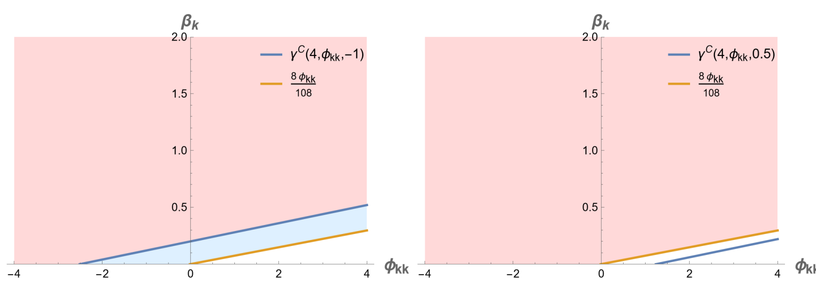
Proposition 17 (The sign of z_k^{C}). *Suppose that $N \geq 2$ and for each $k \in \{b, s\}$, (ϕ_{kk}, β_k) satisfies (3.23). Then we can further partition the domain specified in (3.23) into two regions, which cluster the sign of z_k^C as long as a local condition on the cross-side externalities hold:*

(i) If $\beta_k > \gamma^C(N, \phi_{kk}, u_k^0)$, then there exists $\varepsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\varepsilon(0) \subset \mathbb{R}^2$, $z_k^C < 0$.

(ii) $\beta_k < \gamma^C(N, \phi_{kk}, u_k^0)$, then there exists $\varepsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\varepsilon(0) \subset \mathbb{R}^2$, $z_k^C > 0$.

The interpretation of Proposition 17 is very similar to that of Proposition 13. When the user's heterogeneity in tastes is small (i.e., $\beta_k < \gamma^{\mathrm{C}}(N, \phi_{kk}, u_k^0)$), then users receive $z_k^{\mathrm{C}} > 0$. On the other hand, if β_k is large (i.e., $\beta_k > \gamma^{\mathrm{C}}(N, \phi_{kk}, u_k^0)$), users receive $z_k^{\mathrm{C}} < 0$. Figure 3.4 demonstrates the regions described in Proposition 17 for two different values of u_k^0.

Figure 3.4: Classification of the sign of z_k^{C} based on (ϕ_{kk}, β_k), $k \in \{b, s\}$, according to Proposition 17, where $N = 4$ and $u_k^0 = -1$ (left) or $u_k^0 = 0.5$ (right). The red and blue regions correspond to negative and positive z_k^{C}, respectively.



Corollary 3 ($\gamma^C(N, \phi_{kk}, u_k^0)$ vs $\gamma(N, \phi_{kk}, u_k^0)$). *If $N \geq 2$ and $\gamma(N, \phi_{kk}, u_k^0) \geq 0$, then $\gamma(N, \phi_{kk}, u_k^0) \geq \gamma^C(N, \phi_{kk}, u_k^0)$.*

By Corollary 3, in order to have a positive normalized net deterministic utility z_k in CE, the size of the user's heterogeneity in tastes must be smaller than in CNE. Moreover, in CE, we also identify a critical threshold for the outside utility such that above this threshold, the condition of (ii) in Proposition 17 is not feasible.

Corollary 4 (The sign of z_k^{C} for large values of u_k^0)**.** *Case (ii) in Proposition 17 is not feasible if* $u_k^0 \geq \tilde{u}_k^C(N, \phi_{kk})$*, where* $\tilde{u}_k^C(N, \phi_{kk})$ *is the critical threshold for the outside utility and it is defined in (B.125) in the Appendix Appendix B.*

Economic Outputs in Competitive vs. Collusive Markets

The following proposition compares the normalized net deterministic utility, market participation and prices in competitive and collusive markets.

Proposition 18 (Competition vs Collusion Outputs)**.** *Suppose that* $N \geq 2$ *and for each* $k \in \{b, s\}$*,* (ϕ_{kk}, β_k) *satisfies* (3.19)*. Then, there exists* $\epsilon > 0$ *such that for any* $\varphi_1 = (\phi_{bs}, \phi_{sb}) \in B_\epsilon(0) \subset \mathbb{R}^2$*, in equilibrium:*

(i) the normalized net deterministic utility for users on side k *is bigger under competition than under collusion (i.e.,* $z_k^* > z_k^C$*);*

(ii) the market participation is bigger under competition than under collusion (i.e., $Nx_k^* > Nx_k^C$*);*

(iii) the price charged on side k *of the market is smaller under competition than under collusion (i.e.,* $p_k^* < p_k^C$*).*

Part (i) of the above proposition agrees with the standard collusion literature (see, e.g., Bishop (1960), Varian (1989), Brander and Spencer (1985) among others) in which users receive the lowest normalized net deterministic utility under collusion. Part (ii) is a direct corollary of part (i). Indeed, Proposition 9 implies that $\omega(\cdot)$ is monotonically increasing. Thus, combining (3.17), (3.22) and part (i) of Proposition 18 leads to part (ii) as follows:

$Nx_k^* = N\omega(z_k^*) > N\omega(z_k^{\mathrm{C}}) = Nx_k^{\mathrm{C}}$. To explain part (iii), we use (3.12) and decompose the difference between collusion and competition prices as follows:

$$\boldsymbol{p}^{\mathrm{C}} - \boldsymbol{p}^* = \Phi(\boldsymbol{x}^{\mathrm{C}} - \boldsymbol{x}^*) + \boldsymbol{\beta}(\boldsymbol{z}^* - \boldsymbol{z}^{\mathrm{C}}). \tag{3.32}$$

competition. A careful combination of this formula with parts (i) and (ii) of Proposition 18, the assumption $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$ (or for simplicity $(\phi_{bs}, \phi_{sb}) = 0$) and the observation that if $\phi_{kk} \geq 0$ for $k \in \{b, s\}$ then $\beta_k > f(N)\phi_{kk}$ (see (3.19)) leads to part (iii). We thus note that the above detailed analysis regarding the deterministic net utility is valuable for deriving broader economic implications.

Our results can be compared to other ones on platform collusion. Dewenter et al. (2011) study collusion and competition following the idiosyncrasies of a newspapers market with two firms. They find that for small cross-side network externalities the collusive price is higher than the competitive price. We generalize this result by incorporating the outside option utility, u_k^0, the within-side network externalities, ϕ_{kk}, and by having N horizontally differentiated platforms. Part (iii) of our result also has the same conclusion as Cohen and Zhang (2022), who in the context of ride-sharing services (e.g., *Uber* and *Lyft*), show that under collusion, riders pay a larger price and workers receive a lower wage than under competition (note that the wage in their model is a negative price in our model). Nevertheless, Cohen and Zhang (2022) assume a different model for the user's utility function, which is tailored for their specific setting of prices and wages.

3.5 The Effects of Increasing Competition on the CNE

We study how increasing competition (i.e., increasing N) affects four CNE quantities: price, market participation, consumer surplus, and profit. We first establish sufficient conditions for the derivative $\partial p_k^*/\partial N$ to be either positive or negative (see Proposition 19).

We thus specify regions where competition can lead to increasing or decreasing prices. We also establish sufficient conditions for $\partial(Nx_k^*)/\partial N$ to be positive and consequently for increasing market participation under competition (see Proposition 20). We further formulate sufficient conditions to have increasing and decreasing consumer surplus, i.e., to have positive and negative $\partial(CS_k^*)/\partial N$ (see Proposition 21). Finally, we establish sufficient conditions for the derivative $\partial\pi_k^*/\partial N$ to be either positive or negative (see Proposition 22). That is, we specify regions where competition can lead to increasing or decreasing profits on side k of the market.

The effect of competition on prices. We study the sign of $\partial p_k^*/\partial N$. We first clarify the difficulty in estimating the latter derivative. In view of (3.17), the equilibrium vector price is $\boldsymbol{p}^* = H(\boldsymbol{z}^*)\Omega(\boldsymbol{z}^*)$. As shown in (3.14) and (3.16), the matrix H and the vector Ω directly depend on N. However, (3.13) and the definitions of H and Ω imply that $\boldsymbol{z}^*$ is an implicit function of N. It is thus hard to determine the sign of $\partial p_k^*/\partial N$. Nevertheless, when the cross-side externalities are sufficiently small, the following proposition establishes sufficient conditions to determine the sign of $\partial p_k^*/\partial N$. It uses the functions $g_p(N)$ and $f_p(N)$ defined in (B.141) and (B.147), respectively, of Appendix Appendix B. We note that $g_p(N)$ and $f_p(N)$ approach 0 and 1, respectively, as $N \to \infty$.

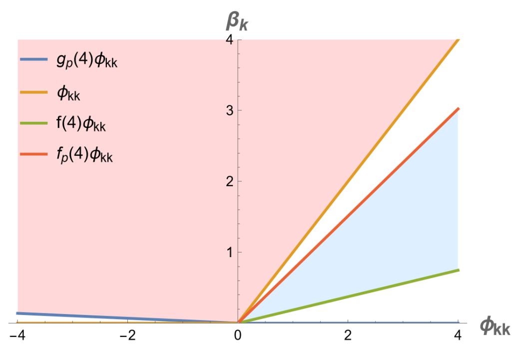
Proposition 19 (Regions where competition decreases/increases prices)**.** *The effect of competition on the change of prices can be clustered into the following two regions:*

(i) Assume that $N \geq 2$ *and either* $(\phi_{kk} \leq 0$ *and* $\beta_k > g_p(N)\phi_{kk})$ *or* $(\phi_{kk} > 0$ *and* $\beta_k > \phi_{kk})$*. Then, there exists* $\epsilon > 0$ *such that for any* $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$*,* $\partial p_k^*/\partial N < 0$*.*

(ii) Assume that $N \geq 3$*,* $\phi_{kk} > 0$ *and* $f(N)\,\phi_{kk} < \beta_k < f_p(N)\,\phi_{kk}$*. Then, there exists* $\epsilon > 0$ *such that for any* $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$*,* $\partial p_k^*/\partial N > 0$*.*

The first part of Proposition 19 agrees with traditional results, where a sufficiently large user's heterogeneity in tastes implies the decrease of the equilibrium prices with the increase of competition, i.e., $\partial p_k^*/\partial N < 0$ (see Anderson and De Palma (1992)). On the other hand, the second part of Proposition 19 agrees with a recent and less conventional result, where positivity of the within-side externalities, ϕ_{kk}, and sufficiently small user's heterogeneity in tastes, β_k imply the increase of prices with the increase of competition, i.e., $\partial p_k^*/\partial N > 0$ (see, Tan and Zhou (2021)). The proposition carefully quantifies the thresholds on the user's heterogeneity in tastes that yield different signs of $\partial p_k^*/\partial N$. The resulting regions are demonstrated below in Figure 3.5 for $N = 4$. The two regions are subsets of the region specified in (3.19) and when $N \to \infty$ the union of the former two regions approaches the latter region (because $g_p(N) \to 0$ and $f_p(N) \to 1$ as $N \to \infty$).

Figure 3.5: Classification of the sign of $\partial p_k^*/\partial N$ based on (ϕ_{kk}, β_k), $k \in \{b, s\}$, according to Proposition 19, where $N = 4$. The red and blue regions correspond to negative and positive $\partial p_k^*/\partial N$, respectively.



The effect of competition on market participation. The equilibrium market participation on side k of the market is given by Nx_k^*, where x_k^* is given by (3.16). The following proposition provides sufficient conditions for positive $\partial(Nx_k^*)/\partial N$. It uses the following quantity:

$$g_x(N) := \frac{(2N^2 - 2N + 1)}{N\,(N^2 - N + 1)}. \tag{3.33}$$

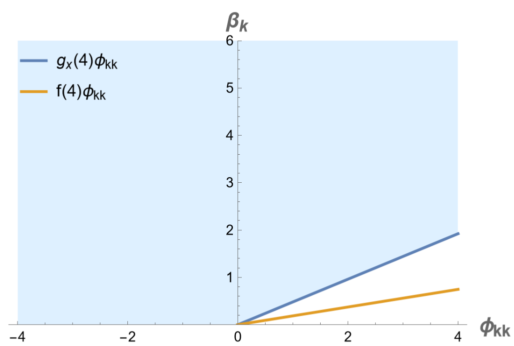
Proposition 20 (Competition increases market participation). *If* $N \geq 2$ *and*

$$\text{either } (\phi_{kk} \leq 0 \text{ and } \beta_k > 0) \text{ or } (\phi_{kk} > 0 \text{ and } \beta_k > g_x(N)\phi_{kk}),$$

then there exists $\epsilon > 0$ *such that for any* $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$*,* $\partial(Nx_k^*)/\partial N > 0$*.*

Most models of platform competition leave out the analysis of the outside utility option. By doing so, they assume full market coverage,[17] and thus cannot study the effect of competition on market participation. Proposition 20 fills this gap and its region of positive $\partial(Nx_k^*)/\partial N$ is demonstrated below in Figure 3.6 when $N = 4$. We note that the region described by Proposition 19 part (ii) intersects with the region described by Proposition 20. Thus, when the within-side externalities are sufficiently large (relative to the user's heterogeneity in tastes) then both prices and market participation increase with competition. At last, we note that the region described by Proposition 20 is a subset of the region described by (3.19) and they coincide as $N \to \infty$.

Figure 3.6: Demonstration of the region of positive $\partial(Nx_k^*)/\partial N$ based on (ϕ_{kk}, β_k), $k \in \{b, s\}$ (in blue), according to Proposition 20, where $N = 4$.



The effect of competition on consumer surplus. The equilibrium consumer surplus on side k of the market, CS_k^*, is defined above in (3.30). The following proposition provides sufficient conditions to determine the sign of $\partial(CS_k^*)/\partial N$. It uses the following quantities:

$$g_{CS}(N) := \frac{2N^3 - N + 1}{N^2(N^2 - N + 2)} \tag{3.34}$$

[17] Corollary 2 shows that full market coverage occurs when $N \to \infty$, however, when the number of platforms is finite, we find this assumption unrealistic.

and $f_{CS}(N)$ which is given by (B.169) in Appendix Appendix B.

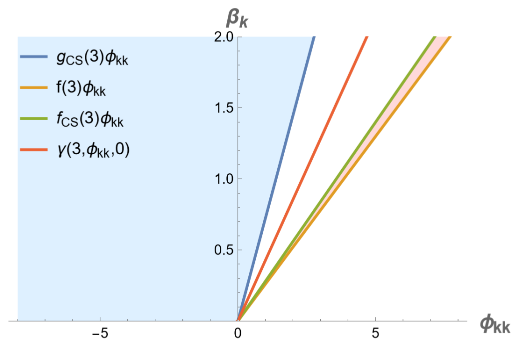
Proposition 21 (Regions where competition decreases/increases consumer surplus)**.** *The effect of competition on the change of consumer surplus can be clustered into the following two regions:*

(i) If $N \geq 2$ and either ($\phi_{kk} \leq 0$ and $\beta_k > 0$) or ($\phi_{kk} > 0$ and $\beta_k > g_{CS}(N)\,\phi_{kk}$), then there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, $\partial(CS_k^*)/\partial N > 0$.

(ii) If $N \geq 7$, $\phi_{kk} > 0$, $f(N)\,\phi_{kk} < \beta_k < \min\{f_{CS}(N)\,\phi_{kk}, \gamma(N, \phi_{kk}, u_k^0)\}$ and $z_k^* < \frac{1}{5}\ln 2$, then there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, $\partial(CS_k^*)/\partial N < 0$.

Part (i) of Proposition 21 agrees with traditional results, where consumer surplus increases with increased competition. For example, Hsu and Wang (2005) consider the Bertrand competition model with substitute goods and show that competition increases consumer surplus. The region in part (i) of Proposition 21 has small cross-side externalities and its within-side externalities are either negative or positive and small with respect to the user's taste heterogeneity. Part (ii), on the other hand, shows sufficient conditions for decreasing consumer surplus with increased competition. This result agrees with a result from Tan and Zhou (2021), where in markets that are relatively concentrated with a few platforms, consumer surplus decreases as competition increases. Moreover, in the asymptotic regime as N goes to infinity, the region in part (ii) disappears (because $g_{CS}(N) \to 0$ and $f_{CS}(N) \to 0$ as $N \to \infty$) and such behavior is also observed in Tan and Zhou (2021). Note that the region in Part (ii) of Proposition 21 has positive within-side externalities, small user's heterogeneity in tastes relative to the within-side externalities, positive but small normalized net deterministic utility relative to the number of platforms, and small cross-side externalities. Figure 3.7 demonstrates the resulting regions (i) and (ii) when $N = 4$, while excluding the condition involving z_k^*.

Figure 3.7: Classification of the sign of $\partial(CS_k^*)/\partial N$ based on (ϕ_{kk}, β_k), $k \in \{b, s\}$, for $N = 4$ and $u_k^0 = 0$. The blue and red regions correspond to positive and negative $\partial(CS_k^*)/\partial N$, respectively. For the red region we did not include the bound on z_k^*, but we still demonstrate a restricted region.



The effect of competition on profits. The equilibrium profit quantity, π^*, is given by

$$\pi^* := \sum_{k \in \{b,s\}} p_k^* x_k^*. \tag{3.35}$$

For each $k \in \{b, s\}$, let $\pi_k^* := p_k^* x_k^*$, the profits on side k of the market. The following proposition provides sufficient conditions to determine the sign of $\partial \pi_k^*/\partial N$. It uses the following condition:

$$\text{either } (\phi_{kk} \leq 0 \text{ and } \beta_k > 0) \text{ or } (\phi_{kk} > 0 \text{ and } \beta_k > g_\pi(N)\, \phi_{kk}), \tag{3.36}$$

where $g_\pi(N) > f(N)$, $f(N)$ is given by (3.18) and $g_\pi(N)$ is given by (B.185) in Appendix Appendix B. It also uses the functions $g_{\pi,z}(N, \phi_{kk}, u_k^0, \beta_k)$ and $f_{\pi,z}(N, \phi_{kk}, u_k^0, \beta_k)$ defined in (B.182) and (B.183) of Appendix Appendix B, respectively.

Proposition 22 (Regions where competition decreases/increases profits on side k)**.** *Assume that $N \geq 2$ and for each $k \in \{b, s\}$, (ϕ_{kk}, β_k) satisfies (3.19). The effect of competition on the change of profits on side k of the market can be clustered into the following two regions of z_k^*:*

(i) If $z_k^* < g_{\pi,z}(N, \phi_{kk}, u_k^0, \beta_k)$*, then there exists* $\epsilon > 0$ *such that for any* $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$*,* $\partial \pi_k^*/\partial N < 0$*.*

(ii) If $z_k^* > f_{\pi,z}(N, \phi_{kk}, u_k^0, \beta_k)$ *and (3.36) is satisfied, then there exists* $\epsilon > 0$ *such that for any* $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$*,* $\partial \pi_k^*/\partial N > 0$*.*

Part (i) of Proposition 22 shows that in markets where the normalized net deterministic utility from joining the market is relatively small, the increased competition decreases profits. In other words, when the incentive to join the market in equilibrium, z_k^*, is small enough, more platforms joining the market reduce the pie for all of the competing platforms. A more interesting result appears in part (ii) when the incentives to join the market are high enough (relatively large z_k^*) and thus the increased competition increases profits. This observation aligns with traditional results in the platform's literature (see Tan and Zhou (2021)) where the effect of network externalities can reverse the usual link between competition and firm profit (i.e., profits can increase with competition).

3.6 Economic and Policy discussion

We examine the economic implications of some of the results presented in Sections 3.4 and 3.5. We first discuss how increases in outside option utility and competition influence equilibrium prices and consumer surplus. These findings may inform policy discussions aimed at improving consumer outcomes and market efficiency. We also interpret our mathematical result comparing collusion and competition under small cross-side externalities. For concreteness, we focus on the dating app market, as motivated in the introduction.

Scenarios in which an increased outside option or greater competition leads to lower prices and higher consumer surplus: Both parts (i) of Propositions 14 and 16 suggest that, under some conditions, such as relatively high heterogeneity, increasing the value of the outside option decreases prices and increases consumer surplus. In the setting of popular dating apps that attract a heterogeneous population—such as Tinder, Bumble, and Hinge—increased preference for traditional partner-finding methods leads to the reduction of dating app prices and the increase in consumer surplus. This suggests that apps should adjust pricing strategies, possibly by reducing prices or enhancing sign-up services. Similarly, both parts (i) of Propositions 19 and 21 imply that, under some conditions, such as relatively high heterogeneity, increasing competition decreases prices and increases consumer surplus. This finding is well-known in the traditional single-sided competition literature (see, e.g., Tirole (1988) and Anderson and De Palma (1992)). In summary, under certain conditions—particularly high heterogeneity—our findings suggest two regulatory mechanisms to decrease prices and increase consumer surplus: enhancing the value of outside options or incentivizing competition.

Scenarios in which an increased outside option or greater competition leads to higher prices and lower consumer surplus: Both parts (ii) of Propositions 14 and 16 suggest that under different conditions, such as relatively low heterogeneity, an increased outside option raises prices and reduces consumer surplus. These results may be exemplified by dating apps that target specific demographics or niches where users are often homogeneous in their preferences. For example, apps like The League and JDate target more homogeneous segments of the population, and consequently, they can charge higher prices. Therefore, in these apps, users are less sensitive to outside options. Moreover, if subscribers are loyal at a sufficiently high outside option utility, there is no incentive to reduce prices even when this utility increases. Similarly, both parts (ii) of Propositions 19 and 21 imply that, under some conditions, such as relatively low heterogeneity, increasing competition leads to higher prices and lower consumer surplus. The intuition follows from (3.30). We first note from this equation that consumer surplus is inversely related to price, so we focus on the former. Additionally, we observe that, for a fixed price, consumer surplus increases with (i) the expected maximum user idiosyncrasy, and (ii) the size of the network externalities. In homogeneous populations, the expected maximum user idiosyncrasy is relatively small, which makes network externalities more pronounced. In this scenario, fewer platforms can amplify network effects more effectively (e.g., instead of having many alternatives to The League or JDate), making them more attractive. As a result, even with increased competition or a higher outside option, users may gravitate toward a smaller number of large platforms to maximize the benefits of network externalities. This dynamic allows these platforms to maintain or increase prices, while consumer surplus remain stagnant or decreases.

Population heterogeneity matters for policy: The discussion above suggests that in markets like those described in this paper, regulators should carefully assess the level of population heterogeneity when aiming to improve consumer surplus and reduce equilibrium prices. This is because the same policy can have varying effects depending on the degree of heterogeneity. Specifically, when population heterogeneity is sufficiently high, policies that either promote competition (e.g., reducing entry barriers and enforcing antitrust laws) or improve the outside option (e.g., enhancing public spaces like parks, libraries, and cultural centers) tend to lower equilibrium prices and increase consumer surplus. Conversely, when population heterogeneity is sufficiently low, policies that limit competition (e.g., supporting a dominant platform) or restrict the outside option (e.g., subsidizing part of the cost for some consumers) can help maintain or reduce prices while preserving or increasing consumer surplus.

Collusion under small cross-side externalities: Proposition 18 shows that in cases of small cross-side externalities, collusion (in comparison to competition) results in decreased normalized net deterministic utilities, reduced market participation and increased price, on both sides of the market. This is intuitive since when the cross-side externalities are sufficiently small—meaning users derive limited benefit from the presence of users on the opposite side—competing platforms have strong incentives to lower prices and attract users. In contrast, colluding platforms internalize each other's pricing decisions and reduce competition, enabling them to raise prices on both sides. This further results in higher net deterministic utilities and greater overall participation for the competition case versus the colluding one. The collusive outcome resembles classic monopoly pricing: platforms extract more surplus at the expense of user welfare, resulting in higher prices and lower market participation compared to the competitive case. In the dating app market, cross-side externalities capture the value one side (e.g., men) derives from a larger presence of the other side (e.g., women) on a given platform. These externalities are typically lower in large-scale casual apps like Tinder, Badoo, and Facebook Dating, where user pools are already extensive and the marginal value of new users is diminished. While Proposition 18 is difficult to verify empirically, we illustrate its logic with a speculative example. During the 2013-2017 period of increasing competition among casual dating apps like Tinder, OkCupid, and Plenty of Fish (POF), prices were lower, user utility was higher, and market participation was larger. In contrast, we hypothesize that the dating app market has shifted in recent years toward reduced competition and arguably increased collusion. Match Group has gained a dominant position through acquisitions of major platforms such as Tinder, POF, OkCupid and Hinge (Gilbert, 2019). In parallel, the adoption of AI-based pricing strategies raises questions about the potential for tacit coordination (Chica et al., 2024a). During this period, rising prices have become evident. Moreover, features that were once free are increasingly placed behind paywalls. This results in lower utility for price-sensitive users and may limit participation, despite overall market growth.

Lastly, we note that the impact of increasing or decreasing competition appears both in Proposition 18, where competition is compared to collusion in an extreme case, and in Propositions 19 and 21, where competition changes by either increasing or decreasing the number of platforms in the market.

3.7 Conclusions

We provided a realistic framework for platform competition and collusion with an outside option. Among our many results, we emphasize the following ones:

1. We showed that when the cross-side externalities are sufficiently small, the normalized net deterministic utilities and market participation are smaller in collusion than competition, and the prices on both sides of the market are bigger in collusion than competition.

2. Depending on the size of the user's heterogeneity in tastes, incorporating an outside option may increase or decrease the equilibrium price and consumer surplus w.r.t. to the no outside option model. In particular, a model of platform competition that omits the outside option will either overestimate or underestimate the true equilibrium price.

3. Depending on the size of the user's heterogeneity in tastes, the number of platforms and the size of network externalities, we also demonstrated when different quantities either decrease or increase with increased competition.[18]

While the paper uses lengthy mathematical derivation, a basic and fundamental idea is demonstrated in (3.32). This equation decomposes the price gap between collusion and competition into two forces: reduced network benefits from lower participation, and lower user utility under collusion. Together, these explain why prices are higher in the collusive regime.

[18] In particular, when the number of platforms increases, prices decrease if the user's heterogeneity is relatively large compared to the within-side externalities, and increase if there are at least three platforms and the user's heterogeneity is relatively small compared to the within-side externalities; market participation always increases; consumer surplus increases if the user's heterogeneity is relatively large compared to the within-side externalities, and decreases if there are at least three platforms, the user's heterogeneity is relatively small compared to the within-side externalities, and the net deterministic utility is small relative to the number of platforms; and profits decrease if the net normalized deterministic utility is small enough and increase if the net normalized deterministic utility is large enough.

There are many open directions for future research. In particular, it would be interesting to extend our model to incorporate the following features: (i) a multi-homing option, i.e., allowing users to join more than one platform; and (ii) platform asymmetries, i.e., allowing for different marginal costs of serving users. Incorporating multi-homing would require introducing an additional decision margin for users, potentially following the frameworks in Chica et al. (2021) or Teh et al. (2023). For the case of platform asymmetries, one could modify problems (3.4) and (3.5) by introducing a marginal cost $c_i > 0$ for each platform $i \in \{1, \ldots, n\}$. Exploring these extensions would likely require a combination of numerical methods and further simplifying assumptions. We view these as promising directions for future work that can build on the foundation laid by the present analysis. Another direction we are currently exploring is the use of our models as an economic framework for analyzing how reinforcement learning algorithms for platform pricing affect equilibrium outcomes. Our models help us assess whether network externalities mitigate or exacerbate the degree of collusion that AI-driven platforms may achieve (Chica et al., 2024a).

Chapter 4

Artificial Intelligence and Algorithmic Price Collusion in Two-sided Markets

4.1 Introduction

Algorithmic price collusion occurs when economic agents set prices using artificial intelligence (AI) algorithms. Through repeated interactions, these agents learn that tacit collusion is optimal, as noted by Calvano et al. (2020a).[1] Economists and antitrust authorities have expressed significant concerns about this form of collusion. The OECD (2017) specifically warned that pricing algorithms could learn to collude through tacit coordination. Assad et al. (2024) suggested that algorithmic pricing in Germany's retail gasoline market increased price margins by approximately 15%. U.S. Senator Amy Klobuchar introduced S.3686, the Preventing Algorithmic Collusion Act of 2024, to curb anticompetitive behavior through algorithmic pricing using nonpublic competitor data.

Recent experiments (Calvano et al. (2020b), Klein (2021)) have demonstrated that collusion can be achieved in Bertrand and Stackelberg competition models by simulated economic agents using Q-learning, a benchmark reinforcement learning algorithm. Building on these findings, this study experimentally investigates AI-driven platforms using Q-

[1] Tacit collusion happens when firms coordinate behavior without explicit communication. Du and Tanriverdi (2023); Bertomeu et al. (2021) have documented evidence of this in the U.S. multihospital system and automotive industry.

learning in a repeated two-sided platform competition game. We show how AI agents facilitate more collusion than Bertrand competition. Our focus is particularly on the impact of network externalities on collusion.

In our model of repeated two-sided platform competition, multiple horizontally differentiated platforms compete to serve buyers and sellers, collectively referred to as users. These platforms repeatedly interact and independently choose prices using Q-learning, with the last period price as the state variable. This implies that platforms have bounded memory and employ one-memory strategies (Barlo et al., 2009). In each repetition, users can choose to join one of the platforms or opt for the outside option. Buyers who join a platform receive network externality benefits proportional to the number of buyers (within-side externalities) and sellers (cross-side externalities) on the same platform. Sellers who join the market receive both types of externalities as well.

Our experiments show that even without network externalities, AI-driven platforms achieve higher collusion levels than those observed in Bertrand competition, as reported by Calvano et al. (2020b). We further conduct experiments to verify that this is due to the higher-dimensional action space, which allows more information exchange. Furthermore, increased network externalities lead to significantly high collusion levels, suggesting AI-driven platforms can leverage these externalities to boost profits. In particular, algorithmic pricing can increase collusion in markets with significant positive within-side externalities (e.g., online/cloud gaming) and positive cross-side externalities (e.g., video streaming, social media). We isolate the effects of network externalities and action space dimensionality to investigate the cause of the high collusion levels. Our results confirm that both the higher-dimensional action space and positive network externalities significantly enhance algorithmic collusion. Our findings also suggest that market participation had only a relatively small impact on profits, while network externalities and Q-learning dynamics played a more direct role in increasing collusion in our two-sided market setting.

Our findings indicate that higher user heterogeneity or greater utility from the outside option generally reduce collusion, except in certain local regions. In contrast, collusion levels typically rise with higher discount rates, especially in the presence of significant network externalities. Notably, tacit collusion remains feasible even at very low discount factors. This contrasts with traditional literature on collusion among firms without AI agents, which suggests that collusion is feasible only at high discount factors (Tirole, 1988; Obara and Zincenko, 2017). Moreover, a supplementary analysis in Appendix Appendix C.1 indicates that platforms tend to learn tacit collusion as part of their equilibrium behavior. Specifically, we show that in a significant portion of cases, algorithms not only choose supracompetitive prices but do so as part of a Nash equilibrium. Similarly, we find that after unilateral price changes by one platform, algorithms frequently revert to the limiting cycle, demonstrating sustained tacit collusive behavior.

Finally, we propose mitigating collusion by incorporating a penalty term into the Q-learning update formula. Moreover, this approach generalizes to any reward-based reinforcement learning method.

Related Literature. There is a growing literature on algorithmic price collusion, with a particular emphasis on using numerical simulations to show that Q-learning results in tacit collusion. Waltman and Kaymak (2008) numerically demonstrated that firms using Q-learning in repeated Cournot oligopoly games produce lower quantities than the competitive Nash equilibrium. Calvano et al. (2020b) showed that Q-learning firms choose high prices in repeated Bertrand games and learn strategies consistent with tacit collusion. Similar work was done by Klein (2021) for repeated Stackelberg games. Assad et al. (2024) is the first work that uses real-life data to show that firms may increase price margins with the adoption of algorithmic pricing. Our work extends the numerical understanding of algorithmic pricing, particularly in two-sided markets with network externalities.

Studies by Johnson et al. (2023) and Brero et al. (2022) on single platforms with AI-driven sellers show how platform-designed rules can promote competition and reduce collusion. Nevertheless, this setting does not apply to ours, where multiple platforms apply AI algorithms.

Our model of repeated two-sided platform competition uses the model in Chica et al. (2023b), which, in turn, builds upon previous models by White and Weyl (2016); Tan and Zhou (2021); Chica et al. (2021). These models analyze network externality effects on equilibrium outputs. In particular, Chica et al. (2023b) provided the dependence of the one-step Nash equilibrium and the one-step collusive equilibrium on market parameters (see Propositions 3.2 and 3.4). Our simulations use the insights of these works to study the impact of these externalities on the collusive levels in repeated games. We are unaware of any work studying the effect of externalities on the equilibrium strategies in a repeated platform game.

Ruhmer (2010) finds that higher cross-side externalities make collusion harder to sustain, when following the model of Armstrong (2006) without AI agents. This is consistent with our numerical results, even though we consider algorithmic pricing and follow the model of Chica et al. (2023b).

Theoretical work in economics on algorithmic price collusion includes Brown and MacKay (2023), which demonstrates that simple pricing algorithms can elevate price levels. Additionally, Arslantas et al. (2024) illustrates how a sophisticated agent can exploit another agent using a naive version of Q-learning, provided the former agent knows the algorithm being used.

Guideline for reading the rest of the paper: Section 4.2 introduces the platform competition framework used in our experiments. Section 4.3 outlines the multi-agent reinforcement learning setup, and Section 4.4 presents extensive numerical experiments. The latter two sections are technical, with a focus on statistical analysis and experimental results. Readers primarily interested in economic interpretations and key intuitions may find Sections 4.5 and 4.6 more relevant, as they discuss the main economic insights and policy implications, respectively. Appendix Appendix C provides sensitivity analyses, examines equilibrium behavior of the Q-agents, and includes additional experiments.

4.2 Review of Our Economic Framework

We introduce the economics framework used in our experiments. Section 4.2.1 presents the baseline platform competition game. Section 4.2.2 extends the latter model to an infinite repeated game.

4.2.1 The Baseline Platform Competition Game

The baseline platform competition game consists of two stages. In stage I, a set of horizontally differentiated platforms strategically choose prices to maximize profits. In stage II, given the prices determined by the platforms, users on each of the two sides of the market choose whether to participate or not and if they participate they also choose which platform to join. The solution concept for the baseline game is backward induction. More specifically, N platforms provide service options for users on two sides of a market, buyers and sellers. Users in these two sides of a market are denoted with $k \in \{b, s\}$, where b and s represent buyers and sellers, respectively. These users have $N + 1$ choices, where $N \geq 2$. They can either opt out of the market by choosing the outside option, or join one of the N horizontally differentiated platforms, each one denoted with $i \in [N] := \{1, \ldots, N\}$. The users on side k opting out of the market receive a deterministic outside option utility $u_k^{(0)} \in \mathbb{R}$. The users on side k joining platform $i \in [N]$ receive a deterministic utility

$$u_k^{(i)} := \phi_k(x_b^{(i)}, x_s^{(i)}) - p_k^{(i)},$$

where $p_k^{(i)}$ is the price paid by the user to access services provided by the platform i; $x_k^{(i)}$ is the total mass of users on side k joining platform i; and

$$\phi_k(x_b^{(i)}, x_s^{(i)}) := \phi_{kb} x_b^{(i)} + \phi_{ks} x_s^{(i)}, \text{ with } \phi_{kb}, \phi_{ks} \in \mathbb{R},$$

is the network externality function that captures the network benefits users enjoy by joining platform i. The network externalities are captured by the following linear transformation

$$(\phi_b(x_b^{(i)}, x_s^{(i)}), \phi_s(x_b^{(i)}, x_s^{(i)}))^T = \Phi \boldsymbol{x}^{(i)}, \text{ where } \Phi = \begin{bmatrix} \phi_{bb} & \phi_{bs} \\ \phi_{sb} & \phi_{ss} \end{bmatrix}. \tag{4.1}$$

To save space, we write $\Phi = [\phi_{bb}, \phi_{bs}; \phi_{sb}, \phi_{ss}]$ when specifying choices of Φ. The endogenous mass of users on each side of the market subscribed to platform i is denoted by $\boldsymbol{x}^{(i)} := (x_b^{(i)}, x_s^{(i)}) \in [0,1]^2$ and the mass of users not participating in the market is denoted by $\boldsymbol{x}^{(0)} := (x_b^{(0)}, x_s^{(0)}) \in [0,1]^2$. Assuming all users have Gumbel-distributed idiosyncratic preferences with parameters (μ_k, β_k), $\mu_k \in \mathbb{R}$ and $\beta_k > 0$,[2] the quantities $x_k^{(i)}$ are determined through a maximization process conducted by users on side k who solve the following equations (see Chica et al. (2023b)):[3]

$$x_k^{(i)} = 1 - \left(1 + \exp\left(u_k^{(i)}/\beta_k - \ln\left(\sum_{j=0,1,\cdots,N, j\neq i} e^{u_k^{(j)}/\beta_k}\right)\right)\right)^{-1} \quad i \in [N] \cup \{0\},\ k \in \{b, s\}. \tag{4.2}$$

The platforms incorporate (4.2) into their profit maximization problem as follows, where $\pi^{(i)}$ denotes the profit of platform i and Π_{tot} denotes the total profits of N colluding platforms:

(i) when competing, each platform i solves

$$\max_{\{p_b^{(i)}, p_s^{(i)}\}} \pi^{(i)}(p_b^{(i)}, p_s^{(i)}), \text{ where } \pi^{(i)}(p_b^{(i)}, p_s^{(i)}) := x_b^{(i)} p_b^{(i)} + x_s^{(i)} p_s^{(i)}; \tag{4.3}$$

[2] The Gumbel-distribution parameter β_k measures the standard deviation of ϵ_k^i and it captures the degree of heterogeneity in users' tastes. Unlike β_k, μ_k does not affect the equilibrium output of the model (see (4.2)).

[3] We note that the model presented here is equivalent to a model in which users have Logistic-distributed preferences for the platforms, and the outside option utility is deterministic.

(ii) when colluding, all platforms jointly solve

$$\max_{p_b, p_s} \Pi_{\text{tot}}(p_b, p_s), \text{ where } \Pi_{\text{tot}}(p_b, p_s) := \sum_{i=1}^{N} \left(x_b^{(i)} p_b + x_s^{(i)} p_s \right). \tag{4.4}$$

The Nash equilibrium associated with (4.3) is referred to as the *competitive Nash Equilibrium* (CNE) and the corresponding equilibrium quantities are denoted by $p_k^{(i),*}$ and $x_k^{(i),*}$ for $i \in [N] \cup \{0\}$ and $k \in \{b, s\}$. The maximizer of (4.4) is called the *collusive equilibrium* (CE) and the corresponding equilibrium quantities are denoted by $p_k^{(i),\text{C}}$ and $x_k^{(i),\text{C}}$, $i \in [N] \cup \{0\}$, $k \in \{b, s\}$. In the symmetric equilibrium, $p_k^{(i),*} = p_k^*$, $p_k^{(i),\text{C}} = p_k^{\text{C}}$, $x_k^{(i),*} = x_k^*$ and $x_k^{(i),\text{C}} = x_k^{\text{C}}$ for all $i \in [N]$. Propositions 3.2 and 3.4 in Chica et al. (2023b) provide first-order conditions for solving p_k^* and p_k^{C}. Similarly, Propositions 3.3 and 3.5 in the same work provides sufficient conditions for the existence and uniqueness of symmetric CNE and CE equilibria. The symmetric CNE and CE individual platform profits are respectively defined by

$$\pi^* := \pi^{(i)}(p_b^*, p_s^*) \text{ and } \pi^{\text{C}} := \frac{\Pi_{tot}(p_b^{\text{C}}, p_s^{\text{C}})}{N}. \tag{4.5}$$

4.2.2 The Infinite Repeated Game

The infinite repeated game consists of a sequence of games, where at time $t \in \mathbb{N} \cup \{0\}$, platforms and users interact following the rules of the baseline platform competition game, introduced in Section 4.2.1, and additional ones. At each time step t, we use the same notations as above, but with a subscript t. We assume that users on all sides are *myopic*, i.e., they make decisions to maximize the utility at current time t by solving (4.2) which depends solely on the current prices observed in the market. We further assume that platforms compete and act strategically and determine the charged prices to maximize the total discounted future rewards at every step t based on the past market states, which we clarify next after introducing some notation and definitions. Given a discounting rate $\delta \in (0, 1)$, we define the total discounted future rewards at time t for platform i by

$$r_t^{(i)} := \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}[\pi_{t+\tau}^{(i)}], \quad \text{where} \quad \pi_{t+\tau}^{(i)} = \sum_{k \in \{b,s\}} x_{t+\tau,k}^{(i)} p_{t+\tau,k}^{(i)} \tag{4.6}$$

and $x_{t+\tau,k}^{(i)}$ is the mass of users on side k joining platform i at time $t+\tau$, and $p_{t+\tau,k}^{(i)}$ is the price that platform i charges on side k at time $t+\tau$. Note that from (4.2), $x_{t+\tau,k}^{(i)}$ is a function of all platforms prices at time $t+\tau$. Furthermore, this observation and (4.3) imply that $\pi_{t+\tau}^{(i)}$ can be written as a function of all platforms prices at time $t+\tau$, that is,

$$\pi_{t+\tau}^{(i)} = \pi^{(i)}(\boldsymbol{p}_{t+\tau}^{(1)}, \boldsymbol{p}_{t+\tau}^{(2)}, \cdots, \boldsymbol{p}_{t+\tau}^{(N)}). \tag{4.7}$$

From the viewpoint of platform i, the policies of all other platforms are unknown, so their present and future prices are considered random variables.[4]

Each platform needs to strategically charge prices in order to maximize the expected total discounted future rewards (4.6). A common method to optimize the expectation of the total discounted future reward is Q-learning, which we introduce in Section 4.3.1.

For $t \in \mathbb{N} \cup \{0\}$, denote by $\boldsymbol{p}_t := (\boldsymbol{p}_t^{(1)}, \ldots, \boldsymbol{p}_t^{(N)})$ the vector of prices chosen by the N platforms at time t, where $\boldsymbol{p}_t^{(i)} := (p_{t,b}^{(i)}, p_{t,s}^{(i)})$, $i \in [N]$. For $L \geq 1$, denoting previous time steps, and $t \geq L$, let

$$\boldsymbol{s}_{t,L} := (\boldsymbol{p}_{t-L}, \boldsymbol{p}_{t-L+1}, \ldots, \boldsymbol{p}_{t-1}) \text{ and } H_{t,L} = \{\boldsymbol{s}_{t,L} \in \mathbb{R}^{2LN}\},$$

where one typically constrains $H_{t,L}$ to be a discrete set (see Section 4.3.2). The problem

[4]For $\tau = 0$, the market share for platform i, $x_{t,k}^{(i)}$, is a random variable since it depends on all the prices charged by all other platforms (see (4.2)). For $\tau > 0$, the future policies of platforms are random variables. Since the future prices depend on the future states, which depend on the platform policies, they are also random variables.

for each platform is to identify a policy

$$\sigma_t^{(i)} : H_{t,L} \longrightarrow \mathbb{R}^2$$

that inputs the current observed state $\boldsymbol{s}_{t,L}$ and outputs the charged price $\boldsymbol{p}_t^{(i)} \equiv (p_{t,b}^{(i)}, p_{t,s}^{(i)})$. During this infinitely repeated game, at each time step t, each platform i updates the policy $\sigma_t^{(i)}$ based on the observed data (the states and rewards) to refine this policy that aims to maximize the expected total discounted future reward. Moreover, at each time step t, each platform i uses the policy $\sigma_t^{(i)}$ to determine the charged prices $\boldsymbol{p}_t^{(i)}$. A particular framework for doing this is discussed in Section 4.3.

4.3 Simulation Framework

We first review the framework of multi-agent reinforcement learning in Section 4.3.1. We then detail our simulation setting in Section 4.3.2, building upon the framework developed in Section 4.3.1.

4.3.1 Preliminaries: Multi-agent Reinforcement Learning

Multi-agent reinforcement learning considers N agents interacting in a dynamic environment. At each time $t \in \mathbb{N}$, each agent $i \in [N]$ observes a state $s_t^{(i)} \in \mathcal{S}$ and takes an action $a_t^{(i)} \in \mathcal{A}$, based on this observed state and following a policy $\sigma_t^{(i)} : \mathcal{S} \longrightarrow \mathcal{A}$, which could be either deterministic or stochastic. Here, $\mathcal{S}$ denotes the state space and $\mathcal{A}$ denotes the action space. Let $\sigma_t = (\sigma_t^{(1)}, \sigma_t^{(2)}, \cdots, \sigma_t^{(N)})$ and $A_t := (a_t^{(1)}, \cdots a_t^{(N)}) = (\sigma_t^{(1)}(s_t^{(1)}), \cdots \sigma_t^{(N)}(s_t^{(N)}))$ denote all policies and actions, respectively, at time t. We denote by $a_t^{(-i)}$, $\boldsymbol{p}_t^{(-i)}$, and $\sigma_t^{(-i)}$ the respective vectors of all actions $a_t^{(j)}$, prices $\boldsymbol{p}_t^{(j)}$, and policies $\sigma_t^{(j)}$with $j \neq i$. The agent collects a reward $\pi_t^{(i)}$, which is a random variable conditioned on the state $s_t^{(i)}$ and actions A_t. The state in the next time, $s_{t+1}^{(i)}$, is a random variable conditioned on the state $s_t^{(i)}$ and the actions A_t taken by all the agents in the current time t. Given a discounting rate $\delta \in (0,1)$, at each time, each agent aims to find a policy in order to maximize the following expectation of the total discounted future reward given all observed states at time t:

$$\sum_{\tau=0}^{\infty} \mathbb{E}_{\pi,s,\sigma}\left[\delta^{\tau}\pi_{t+\tau}^{(i)}(s_{t+\tau}^{(i)}, A_{t+\tau})\right]. \tag{4.8}$$

The expectation is needed due to the randomness in the rewards, the future states, and the future actions of all the agents.

Q-learning is a classic method for finding the policy that maximizes (4.8). It uses the Q-function of agent i at state s given an action a, which is defined by

$$Q^{(i)}(s,a,\sigma^{(i)};\sigma^{(-i)}) := \\ \textstyle\sum_{\tau=0}^{\infty} \mathbb{E}_{\pi,s,\sigma}\left[\delta^{\tau}\pi_{t+\tau}^{(i)}\middle| s_t^{(i)} = \boldsymbol{s}, a_t^{(i)} = a, a_{t+u}^{(i)} = \sigma^{(i)}(s_{t+u}^{(i)}),\ u \geq 1,\ a_{t+v}^{(-i)} = \sigma^{(-i)}(s_{t+v}),\ v \geq 0\right]. \tag{4.9}$$

Note that (4.9) differs from (4.8) by having agent i follow the given action a at time t instead of the policy $\sigma^{(i)}$, whereas in both formulations all other agents at times $t, t+1, \ldots,$ and agent i at times $t+1, t+2, \ldots,$ follow their policies.

We denote an optimal policy for agent i by $\sigma^{(i)*}$, which is hard to find. Q-learning overcomes this difficulty by carefully estimating the solution $Q^{(i)*}(s_t, a_t; \sigma^{(-i)})$ to the following Bellman equation

$$Q^{(i)*}(s_t, a_t; \sigma^{(-i)}) = \mathbb{E}_{\pi}\left[\pi(s_t, A_t)\right] + \delta \max_{a'} \mathbb{E}_{s_{t+1}}\left[Q^{(i)*}(s_{t+1}, a'; \sigma^{(-i)})|A_t\right]. \tag{4.10}$$

It then estimates $\sigma^{(i)*}$ using the following relationship between $\sigma^{(i)*}$ and $Q^{(i)*}(x, a; \sigma^{(-i)})$:

$$\sigma^{(i)*}(s) = \text{argmax}_a Q^{(i)*}(s, a; \sigma^{(-i)}). \tag{4.11}$$

We detail the methods for estimating the Q^*-function in (4.10) in the following section.

4.3.2 The Simulation Setup

We consider a market with two platforms, that is, we set $N = 2$.[5] At time t, each platform $i \in \{1, 2\}$ observes the following state $\boldsymbol{s}_t^{(i)} := \boldsymbol{p}_{t-1}$, which contains prices at the previous step. Platform i determines its prices $\boldsymbol{p}_t^{(i)}$ based on the observed state $\boldsymbol{s}_t^{(i)}$. At each time t, after all platforms have chosen prices $\boldsymbol{p}_t^{(i)}$, they receive the reward $\pi_t^{(i)} = \sum_{k\in\{b,s\}} x_{t,k}^{(i)} p_{t,k}^{(i)}$, where $x_{t,k}^{(i)}$ is solved using (4.2).

To simplify the computation, we allow platforms to choose from a discrete set of M prices. While it is common to expect that $p_k^* < p_k^{\mathrm{C}}$,[6] our model also allows the case $p_k^{\mathrm{C}} < p_k^*$. We further introduce the parameter $\epsilon = 0.1$ so the lowest price is slightly lower than $\min(p_k^*, p_k^{\mathrm{C}})$ and the highest one is slightly higher than $\max(p_k^*, p_k^{\mathrm{C}})$. For each $k \in \{b, s\}$, our set of prices is

$$\mathcal{P}_k := \left\{ p_k^* - \epsilon(p_k^{\mathrm{C}} - p_k^*) + \frac{j}{M-1}(1+2\epsilon)(p_k^{\mathrm{C}} - p_k^*) \middle| j = 0, \cdots, M-1 \right\}. \tag{4.12}$$

Note that the cardinality of the price space $|\mathcal{P}_k|$ is M for both $k = b$ and $k = s$. The overall state space (for both platforms) and the action space for each platform are respectively defined by

$$\mathcal{S} := (\mathcal{P}_b \times \mathcal{P}_s) \times (\mathcal{P}_b \times \mathcal{P}_s) \text{ and } \mathcal{A} := \mathcal{P}_b \times \mathcal{P}_s. \tag{4.13}$$

We note that the size of the state space is $|\mathcal{S}| = M^4$ and the size of the action space is

[5] Since our model assumes two sides of the market, each of the N platforms must choose two prices. At each stage, our simulation estimates $2N$ different prices, and there are M^{2N} possibilities for the vector of prices, where M is the size of the set of price choices available to each platform. To make our simulations feasible, we choose $N = 2$.

[6] Note that Proposition 4.11 in Chica et al. (2023b) provides sufficient conditions to guarantee that $p_k^* < p_k^{\mathrm{C}}$.

$|\mathcal{A}| = M^2$.

Platform policy: We denote the estimation at time t of $Q^{(i)*}(\boldsymbol{s}, \boldsymbol{a}; \sigma^{(-i)})$ by $Q_t^{(i)}(\boldsymbol{s}, \boldsymbol{a})$, where $\boldsymbol{s} \in \mathcal{S}$, $\boldsymbol{a} \in \mathcal{A}$ and $i \in \{1, 2\}$ indexes the platform. Q-learning alternately estimates $Q_t^{(i)}$ and the stochastic policies at time t. We first assume that $Q_t^{(i)}$ is known and show how the platforms determine the stochastic policy at time t. We then explain the Q-learning estimation of $Q_t^{(i)}$. Instead of directly computing the policy as the maximum value in (4.11), Q-learning computes a softmax value using a temperature parameter $\mathcal{T}_t$. For this purpose, at time t and given a state $\boldsymbol{s}_t^{(i)} \in \mathcal{S}$ and the Q^*-function estimate, $Q_t^{(i)}$, the policy of platform i is the Boltzmann probability distribution:

$$P(\boldsymbol{a}_t^{(i)} = \boldsymbol{a} | \boldsymbol{s}_t^{(i)}) = \exp\left(Q_t^{(i)}(\boldsymbol{s}_t^{(i)}, \boldsymbol{a})/\mathcal{T}_t\right) / \sum_{\boldsymbol{a}' \in \mathcal{A}} \exp\left(Q_t^{(i)}(\boldsymbol{s}_t^{(i)}, \boldsymbol{a}')/\mathcal{T}_t\right). \tag{4.14}$$

We remark that all platforms independently determine their prices based on (4.14).[7]

Q-learning estimation: At each time step, after determining the price following (4.14), platform i collects the reward $\pi_t^{(i)}$ defined by (4.7). Next, platform i updates the estimated values of the Q^*-function at the given state $\boldsymbol{s}_t^{(i)}$ and the selected action $\boldsymbol{a}_t^{(i)}$ with a learning rate α as follows:

$$Q_{t+1}^{(i)}(\boldsymbol{s}_t^{(i)}, \boldsymbol{a}_t^{(i)}) := (1-\alpha) Q_t^{(i)}(\boldsymbol{s}_t^{(i)}, \boldsymbol{a}_t^{(i)}) + \alpha \left(\pi_t^{(i)} + \delta \max_{\boldsymbol{a}} Q_t^{(i)}(\boldsymbol{s}_{t+1}^{(i)}, \boldsymbol{a})\right). \tag{4.15}$$

We remark that (4.15) is an approximation of (4.10) (see Watkins and Dayan (1992)).

We initialize the Q^*-function at $\boldsymbol{s} \in \mathcal{S}$ and $\boldsymbol{a} \in \mathcal{A}$ assuming that in all future states platform i charges $\boldsymbol{a}$ and all other platforms charge the prices in $\boldsymbol{s}$. Therefore, for platform i, state $\boldsymbol{s}$, a given action $\boldsymbol{a}$ and the price vector for platform $j \neq i$, which we denote by $\boldsymbol{p}^{(j)}$

[7] As $\mathcal{T}_t$ decreases, (4.14) increasingly focuses on the optimal action based on $Q_t^{(i)}$. When $\mathcal{T}_t \to 0$, the policy randomly selects between the actions that yield the maximal reward $Q_t^{(t)}$ with uniform probabilities. In the simulation, we set $\mathcal{T}_0 = 1000$ to encourage exploration of possible actions, gradually decreasing it towards 0 to exploit optimal actions.

and it is part of the state $\boldsymbol{s}$, the Q^*-function for platform i is initialized by

$$Q_0^{(i)}(\boldsymbol{s}, \boldsymbol{a}) = \sum_{\tau=0}^{\infty} \delta^\tau \pi^{(i)}(\boldsymbol{a}, \boldsymbol{p}^{(j)}) = \frac{\pi^{(i)}(\boldsymbol{a}, \boldsymbol{p}^{(j)})}{1-\delta}. \tag{4.16}$$

Parameter setup: We choose exponentially decaying temperature parameter $\mathcal{T}_t = \mathcal{T}_0 \lambda^t$ with $\mathcal{T}_0 := 1000/(1-\delta)$ and $\lambda = 1 - 10^{-7}$. This choice encourages exploration in the early stages and exploits optimality in the later stages. We choose both the idiosyncratic preference parameters and the outside option utilities to be the same on both sides of the market. Therefore we denote $\beta_k = \beta_b \equiv \beta_s$ and $u_k^{(0)} = u_b^{(0)} \equiv u_s^{(0)}$. We set the learning rate $\alpha = 0.15$, discount rate $\delta = 0.05$, idiosyncratic preferences $\beta_k = 1$, and outside option utility $u_k^{(0)} = -2$ for each $k \in \{b, s\}$. We choose a small value for δ, compared to the choice of the same parameter in Calvano et al. (2020b); Klein (2021), to emphasize that in our setting collusion is already present with a very small discount rate.

Reporting metric: We define the collusive level of platform i at time t as

$$\Delta_t^{(i)} := \frac{\pi_t^{(i)} - \pi^*}{\pi^{\mathrm{C}} - \pi^*}, \tag{4.17}$$

where we recall (see (4.5)) π^* and π^{C} respectively denote the CNE and CE equilibrium profits of the baseline platform competition game. When $\Delta_t^{(i)} = 0$, platform i's reward at time t equals the CNE level, π^*; whereas when $\Delta_t^{(i)} = 1$, it equals the CE level, π^{C}. Each simulation runs $T = 5 \times 10^8$ iterations and we report the overall collusive level in the last $K = 1,000$ steps as follows:

$$\tilde{\Delta} := \frac{1}{KN} \sum_{s=0}^{K-1} \sum_{i=1}^{N} \Delta_{T-s}^{(i)}. \tag{4.18}$$

4.4 Experimental Results

We report extensive numerical experiments using the setup of Section 4.3.2. Section 4.4.1 investigates the general dependence of $\tilde{\Delta}$, defined in (4.18), on the externality matrix Φ. Section 4.4.2 further explores the latter dependence for concrete and useful choices of Φ. Section 4.4.3 studies the dependence of $\tilde{\Delta}$ on the degree of heterogeneity in users' tastes, the outside option utility, and the discount rate. Lastly, Section 4.4.4 explores two interesting scenarios: 1) long-run asymmetric equilibria outperform the symmetric equilibrium, and 2) competition prices are larger than collusion prices for one side of the market. A supplementary analysis in Appendix Appendix C.1 implies that our numerical results are consistent with platforms learning tacit collusion and equilibrium strategies. In particular, we show that in at least 30% of cases (with percentages rising to 50% for different values of Φ), algorithms result in Nash equilibrium behavior. Similarly, in at least 75% of cases (with percentages rising to 95% for different values of Φ), algorithms converge back to the limiting cycle after one unilateral price change by one of the competing platforms.

4.4.1 Dependence of the Collusive Level on the Network Externalities

We applied an additive model to infer the dependence of the collusive level, $\tilde{\Delta}$, on the externality matrix, Φ. We ran 2,500 simulations according to the setting described in Section 4.3.2. For each simulation, we randomly sampled the elements of the externality matrix Φ from independent normalized Gaussians (that is, $\phi_{kl} \sim N(0,1)$ for $k,l \in \{b,s\}$), and recorded the final collusive level, $\tilde{\Delta}$. In order to infer the dependence of $\tilde{\Delta}$ on Φ, we assume the following additive model:

$$\tilde{\Delta}(\Phi) = \Delta_0 + f_{bb}(\phi_{bb}) + f_{ss}(\phi_{ss}) + f_{bs}(\phi_{bs}) + f_{sb}(\phi_{sb}) + f_{bb,ss}(\phi_{bb}, \phi_{ss}) + f_{sb,bs}(\phi_{sb}, \phi_{bs}) + f_{bb,bs}(\phi_{bb}, \phi_{bs}) + f_{bb,sb}(\phi_{bb}, \phi_{sb}) + f_{ss,bs}(\phi_{ss}, \phi_{bs}) + f_{ss,sb}(\phi_{ss}, \phi_{sb}) + \epsilon, \tag{4.19}$$

where Δ_0 is the sample mean of $\tilde{\Delta}$, the next 4 functions (f_{bb}, f_{ss}, f_{bs}, f_{sb}) represent the univariate effects of the elements of Φ on $\tilde{\Delta}$, the last 6 functions ($f_{bb,ss}$, $f_{sb,bs}$, $f_{bb,bs}$, $f_{bb,sb}$, $f_{ss,bs}$, $f_{ss,sb}$) represent the bivariate effects of the elements of Φ on $\tilde{\Delta}$ and ϵ is an error term, encompassing higher-order multivariate effects. Since the equilibrium values π^* and π^C depend on Φ nonlinearly (see Section 4.2.1), the 10 functions, $f_{bb}, \cdots, f_{ss,sb}$, are nonlinear. We thus sequentially fit these functions using XGBoost (Chen and Guestrin, 2016), which is a popular non-parametric, nonlinear fitting method. To reduce the bias of the fitted functions, we alter the order of both the first four functions and the next six functions, during the sequential fitting procedure, and average the collusive level over the different orders. Appendix Appendix C.2 contains more details of implementing XGBoost.

We refer to Δ_0 as the baseline collusive level, whereas $\tilde{\Delta}$ is the collusive level. Our simulations show that Δ_0 is approximately 0.3. Next, we report our estimates for the univariate and bivariate effects of the elements of Φ on $\tilde{\Delta}$.

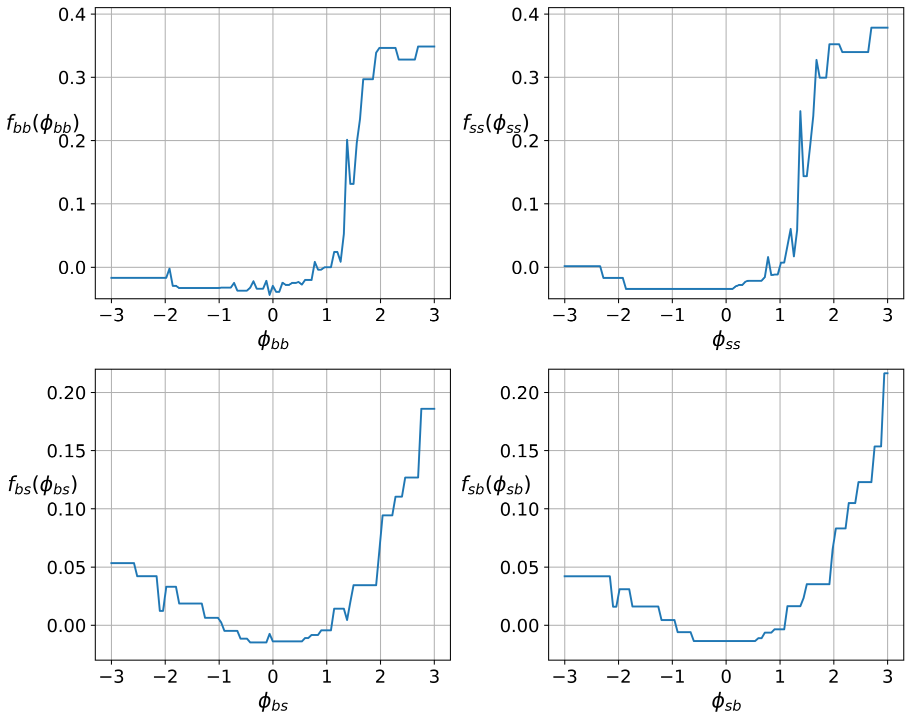


Figure 4.1: Demonstration of the dependence of the four fitted univariate functions on the externalities. Top left: $f_{bb}(\phi_{bb})$; Top right: $f_{ss}(\phi_{ss})$; Bottom left: $f_{bs}(\phi_{bs})$; Bottom right: $f_{sb}(\phi_{sb})$.

Figure 4.1 illustrates the fitted functions f_{bb}, f_{ss}, f_{bs} and f_{sb}, which capture the univariate effect of each entry in the externality matrix. The top two subfigures demonstrate the univariate effect of the within-side externalities, (ϕ_{bb} and ϕ_{ss}). In this case, the collusive level is close to zero when these externalities are less than 1, then increases sharply when these externalities increase from 1 to 2, and it is approximately flat when these externalities are above 2 with a possible increase of the collusive level when the absolute values of the negative externalities increase. We remark that we cannot confidently conclude the latter increase from the current experimental results, but latter experiments in Section 4.4.2 support such an increase, especially when considering lower values of ϕ_{bb} and ϕ_{ss}. The bottom two subfigures demonstrate the univariate effect of the cross-side externalities (ϕ_{bs} and ϕ_{sb}). In this case, the dependence of the collusive level on the externalities is depicted by a J-shape function with a minimum when the externality is around zero. We thus note that in order to minimize the level of the algorithmic collusion, we would need to bound the values of the within-side externalities and the absolute values of the cross-side externalities. In our particular experimental setting, the desired bound is 1. In general, we expect there can be two different upper bounds for the within-side and cross-side externalities and they depend on the chosen parameters, in particular, $\{\delta, \beta_k, u_k^{(0)}\}$.

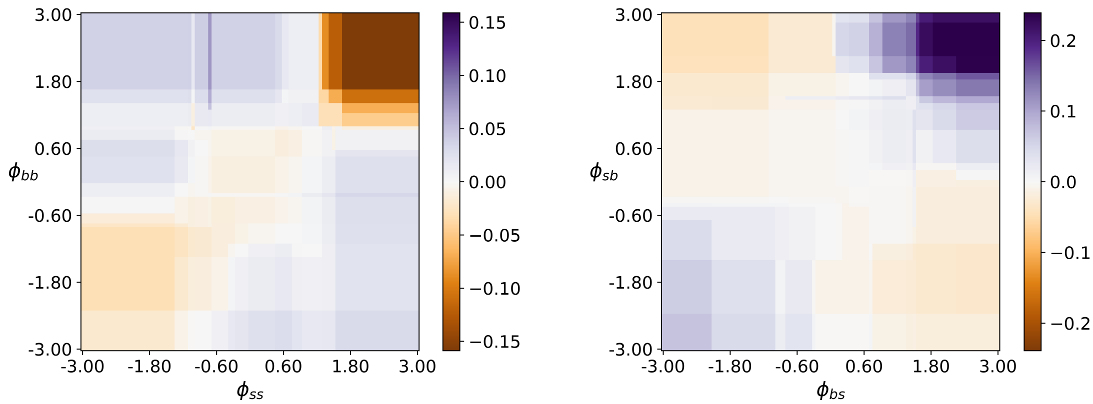


Figure 4.2: Heatmaps of the fitted functions $f_{bb,ss}(\phi_{bb}, \phi_{ss})$ (left) and $f_{sb,bs}(\phi_{bs}, \phi_{sb})$ (right), which capture the bivariate effect between ϕ_{bb} and ϕ_{ss}, and between ϕ_{bs} and ϕ_{sb}, respectively.

Figure 4.2 demonstrates the fitted functions $f_{bb,ss}$ and $f_{sb,bs}$, which capture the bivariate effects on collusion of the main diagonal and off-diagonal elements in the externality matrix. We present the images of these functions as heatmaps over their planar domains. For example, in the left-hand subfigure the domain is described by the within-side externality variables ϕ_{bb} and ϕ_{ss} and the collusive level is depicted by a heatmap, changing from purple (highly positive) to orange (highly negative). The left-hand subfigure implies that when the within-side externalities, ϕ_{bb} and ϕ_{ss}, are both large, they result in the minimal value of the bivariate effect, which is negative. In this regime, the collusive level, resulting from both the univariate and bivariate effects, remains positive (recall that the univariate effect is demonstrated in the top subfigures of Figure 4.1). Similarly, when ϕ_{bb} and ϕ_{ss}

are both sufficiently negative, their bivariate effect reduces the collusive level, albeit by a small amount. We remark that both the univariate and bivariate contributions in this case are rather small and it is hard to predict their combined effect from this experimental result, but another experiment in Section 4.4.2 indicates that they cancel each other. The right-hand side subfigure indicates that the bivariate component of the cross-side externalities reduces the collusive level when these externalities are large in absolute values and have opposite signs. On the other hand, it increases the collusive level when the cross-side externalities have the same sign and have sufficiently large absolute values. The rate of increase is larger when they are both positive.

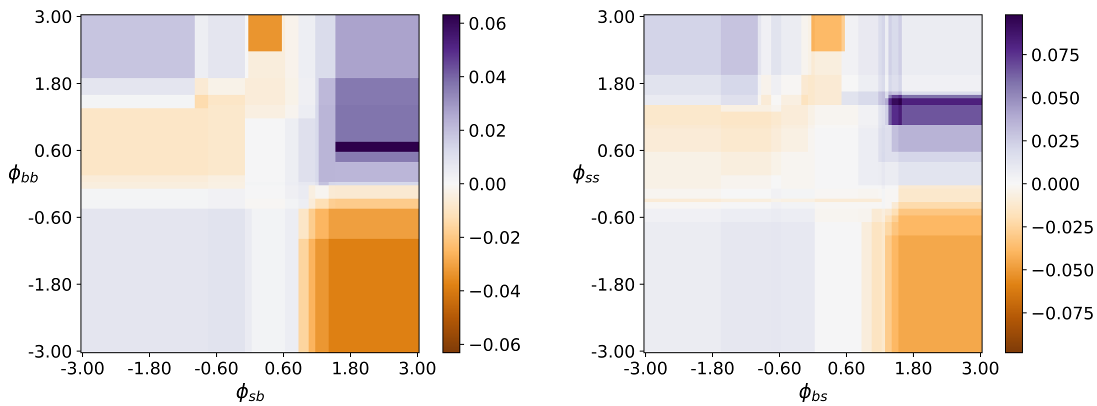


Figure 4.3: Heatmaps of the fitted functions $f_{bb,sb}(\phi_{bb}, \phi_{sb})$ (left) and $f_{ss,bs}(\phi_{ss}, \phi_{bs})$ (right), which capture the bivariate effect between ϕ_{bb} and ϕ_{sb}, and between ϕ_{ss} and ϕ_{bs}, respectively.

Figure 4.3 demonstrates the bivariate effect on collusion when both buyers and sellers benefit from population joining the market on either side b or s (but not both at the same time). That is, it demonstrates the bivariate effect for ϕ_{bb} and ϕ_{sb} when considering side b (left) and the bivariate effect for ϕ_{ss} and ϕ_{bs} when considering side s (right). The two subfigures are very similar and we thus only discuss the left one, with the variables ϕ_{bb} and ϕ_{sb}. The bottom-right corner of this subfigure implies that if ϕ_{sb} is sufficiently large and ϕ_{bb} is sufficiently negative, the bivariate effect on the collusive level is negative. In this regime, the collusive level, resulting from both the univariate and bivariate effects, remains positive (recall that the univariate effect is demonstrated in the top-left and bottom-right subfigures of Figure 4.1). On the other hand, the top-right corner in Figure 4.3 shows that when ϕ_{bb} and ϕ_{sb} are both large the bivariate effect on the collusive level is positive.

Figure 4.4 illustrates the bivariate effect on collusion when either buyers or sellers (but not both at the same time) benefit from population joining the market on sides b or s. That is, it demonstrates the bivariate effect for ϕ_{bb} and ϕ_{bs} when considering only buyers (left subfigure) and the bivariate effect for ϕ_{ss} and ϕ_{sb} when considering only sellers (right subfigure). The two subfigures are very similar and we thus only discuss the left one, with the variables ϕ_{bb} and ϕ_{bs}. We notice that when the ϕ_{bb} is sufficiently large and ϕ_{bs} is sufficiently negative, the bivariate effect on the collusion is negative. In this regime, the collusive level, resulting from both the univariate and bivariate effects, remains positive (the univariate effect is demonstrated in the left subfigures of Figure 4.1). We further notice that when ϕ_{bs} is sufficiently large and ϕ_{bb} is sufficiently negative the bivariate effect on the collusion is also negative, but smaller than the latter one. Similarly, the collusive level, resulting from both the univariate and bivariate effects, remains positive. On the other hand, when both ϕ_{bb} and ϕ_{bs} are sufficiently large, the bivariate effect on the collusive level is positive.

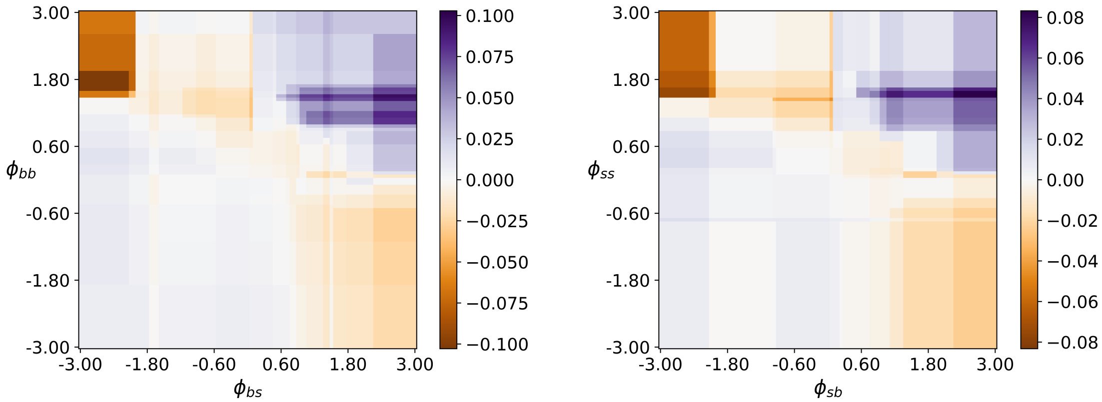


Figure 4.4: Heatmaps of the fitted functions $f_{bb,bs}(\phi_{bb}, \phi_{bs})$ (left) and $f_{ss,sb}(\phi_{ss}, \phi_{sb})$ (right), which capture the bivariate effect between ϕ_{bb} and ϕ_{bs}, and between ϕ_{ss} and ϕ_{sb}, respectively

4.4.2 A Study of the Collusive Level under Special Network Externalities

We assume special parameterizations of the network externality matrices, Φ, and explore the dependence of $\tilde{\Delta}$ on any such Φ. This allows us to track more carefully the dependence of $\tilde{\Delta}$ on Φ in some special settings. For each specific Φ, we ran 100 simulations. Our figures present the dependence of the overall collusive level on the elements of Φ, where their main curves represent the average of the collusive levels from the 100 runs and their shaded areas represent the uncertainty level, which was computed using bootstrapping with a 99% confidence interval.[8]

Figure 4.5 investigates the dependence of the collusive level on the within-side externalities in two controlled settings. In the first setting (left panel) $\Phi = [\phi_{bb}, 0; 0, 0]$, and in the second one (right panel) $\Phi = [\phi_{bb}, 0; 0, \phi_{bb}]$. In both cases $\phi_{bb} \in [-6, 2]$.

[8] Both bootstrapping and visualization are implemented using the `seaborn` Python package.

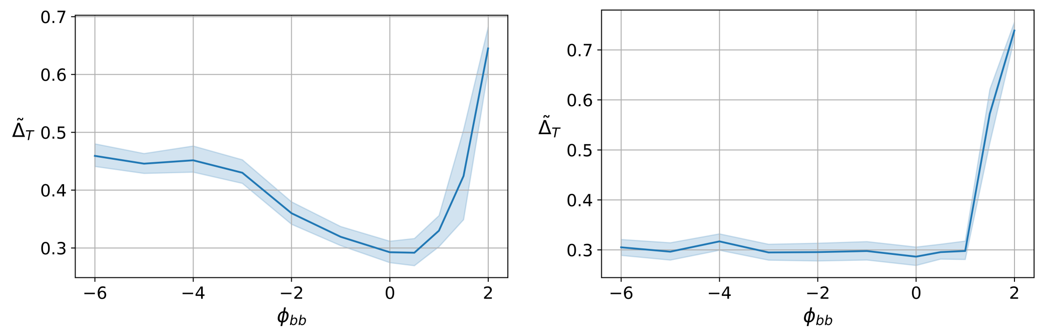


Figure 4.5: Collusive level with varying ϕ_{bb}: $\Phi = [\phi_{bb}, 0; 0, 0]$ (left) and $\Phi = [\phi_{bb}, 0; 0, \phi_{bb}]$ (right).

In the left panel, the collusive level forms a J-shape, where it decreases on $[-6, 0]$ and increases sharply on $[0.5, 2]$. The minimum value of the collusive level is achieved when $\phi_{bb} \approx 0.5$ and it is slightly below the baseline collusive level. Note that this subfigure indicates a similar behavior of the collusive level to its univariate effect shown in the top left panel of Figure 4.1. Indeed, in this case, the collusive level depends on the single variable ϕ_{bb}, so the other univariate and bivariate functions are irrelevant. However, the minimal value of the univariate effect in the top left panel of Figure 4.1 is around zero, since it is separate from the baseline collusive level Δ_0. By adding Δ_0 to this univariate effect, we obtain a function similar to the collusive level described in the left panel of Figure 4.5. We remark that in the experiments of Section 4.4.1, our domain was restricted by the underlying Gaussian model and thus the domain in Figure 4.1 is narrower than that of Figure 4.5.

In the right panel, the collusive level sharply increases when ϕ_{bb} exceeds 1. This behavior can be explained using our previous findings. Indeed, as shown in Figures 4.1 and 4.2, when the within-side externalities, ϕ_{bb} and ϕ_{ss}, are both large, the univariate effect is more significant than the negative bivariate effect, resulting in a significant increase. We also notice that the collusive level remains flat and around Δ_0 when $\phi_{bb} = \phi_{ss}$ falls below 1. This observation also confirms our findings in the previous section. Indeed, Figures 4.1 and 4.2

indicate that when ϕ_{bb} and ϕ_{ss} are both sufficiently negative, the (positive) univariate and (negative) bivariate effects cancel each other.

Figure 4.6 investigates the dependence of the collusive level on the cross-side externalities in two controlled settings. In the first setting (left panel) $\Phi = [0, \phi_{bs}; 0, 0]$, and in the second one (right panel) $\Phi = [0, \phi_{bs}; \phi_{bs}, 0]$. In both cases $\phi_{bs} \in [-2.5, 3]$.

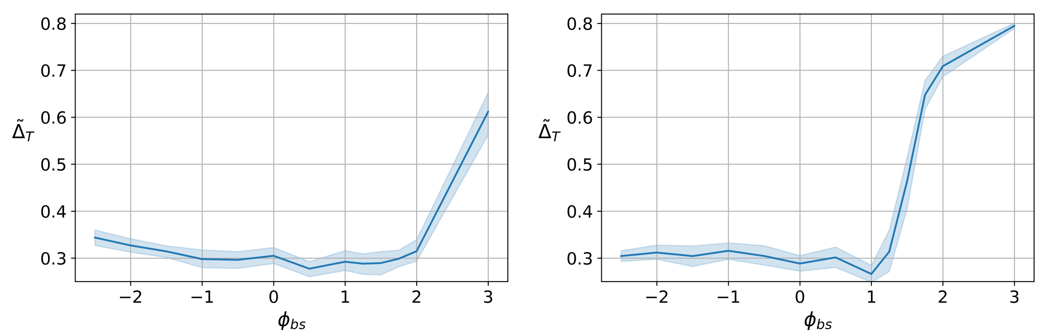


Figure 4.6: Collusive level with varying ϕ_{bs}: $\Phi = [0, \phi_{bs}; 0, 0]$ (left) and $\Phi = [0, \phi_{bs}; \phi_{bs}, 0]$ (right).

In the left panel, the collusive level increases when the cross-side externality exceeds 2, and it slightly decreases when the same externality falls below -1. This observation aligns with our findings in the previous section. Indeed, as shown by the bottom panels in Figure 4.1, the collusive level increases as ϕ_{bs} increases in absolute value with values above 1.

In the right panel, the collusive level increases when the cross-side externalities exceed 1. It has a sharper increase than the one in the left panel. These observations agree with the findings of the previous section. Indeed, Figures 4.1 and 4.2 show that the univariate and bivariate effects of ϕ_{bs} and ϕ_{sb} are both positive when $\phi_{bs} = \phi_{sb} > 1$. Furthermore, Figure 4.2 shows the positive bivariate effect between ϕ_{bs} and ϕ_{sb}, which explains the sharper increase in the right panel. The dependence of the collusive level in the right panel on smaller values of ϕ_{bs}, which are not shown in this figure, is rather unique and thus deferred to Section 4.4.4.

Figure 4.7 investigates the dependence of the collusive level on the bivariate effects between the within- and cross-side externalities in two controlled settings. For simplicity, we fix the cross-side externality and vary the within-side externality. In the first setting (left panel) $\Phi = [\phi_{bb}, 3; 0, 0]$, and in the second one (right panel) $\Phi = [\phi_{bb}, 3; 0, \phi_{bb}]$. In both cases $\phi_{bb} \in [-6, 2]$.

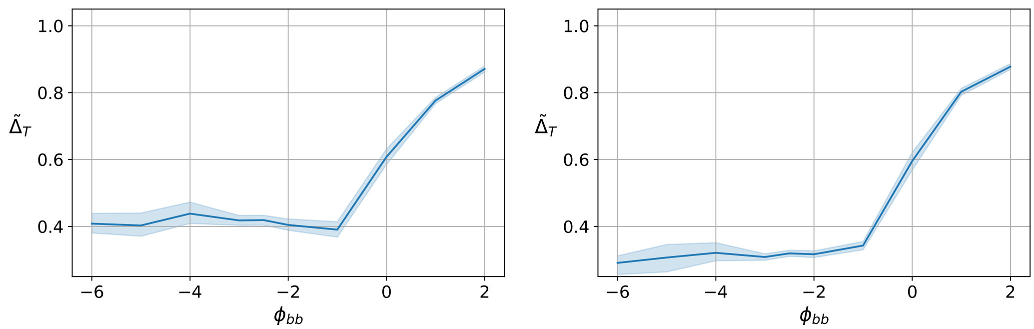


Figure 4.7: Collusive level with varying ϕ_{bb}: $\Phi = [\phi_{bb}, 3; 0, 0]$ (left) and $\Phi = [\phi_{bb}, 3; 0, \phi_{bb}]$ (right).

In the left panel, the collusive level is flat when the within-side externality falls below -1 and increases sharply as it exceeds -1. This increase can be explained when ϕ_{bb} exceeds 0 by the univariate effect for ϕ_{bb} shown in Figure 4.1. On the other hand, the increase in $[-1, 0]$ can be explained by the bivariate effect between ϕ_{bb} and ϕ_{bs}, shown in Figure 4.4. Indeed this bivariate effect increases with respect to ϕ_{bb} in $[-1, 0]$ when $\phi_{bs} = 3.0$.

In the right panel, the collusive level follows a similar pattern as in the left one. This follows from a similar explanation as above, where one should also note that when $\phi_{bs} = 3$, the bivariate effect between ϕ_{ss} and ϕ_{bs} increases with respect to ϕ_{ss} in $[-1, 0]$, as shown in Figure 4.3. Additionally, when $\phi_{bb} < 0$, the collusive level in the right panel is lower than that in the left panel (it is easiest to see this for $\phi_{bb} < -1$). We clarify this observation in view of the findings of Section 4.4.1 as follows. We note that according to the right panel of Figure 4.3 , when $\phi_{bs} = 3$, the bivariate effect between ϕ_{ss} and ϕ_{bs} is negative when $\phi_{ss} < 0$, therefore the collusive level in the right panel is expected to be lower than the collusive level in the left panel when $\phi_{bb} = \phi_{ss} < 0$.

We make some additional remarks comparing Figures 4.5 and 4.7. The left panel in Figure 4.5 shows that the collusive level decreases with respect to ϕ_{bb}, when $\phi_{bb} < 0$ and $\phi_{bs} = 0$. On the other hand, in the left panel of Figure 4.7, the collusive level is flat or increases with respect to ϕ_{bb} when $\phi_{bb} < 0$ and $\phi_{bs} = 3$. This behavior can be explained by the contribution from the bivariate effect between ϕ_{bb} and ϕ_{bs} when ϕ_{bb} is negative. Indeed, this bivariate effect is almost flat with respect to ϕ_{bb} when $\phi_{bs} = 0$, but is increasing with respect to ϕ_{bb} when $\phi_{bs} = 3.0$ (see left panel of Figure 4.4). A similar comparison can be made for the right panels in Figures 4.5 and 4.7, and the explanation similarly follows from the right panel of Figure 4.3 and the left panel of Figure 4.4.

4.4.3 A Study of the Collusive Level under Special Market Parameters

We explore the dependence of the collusive level on the market parameters β_k, $u_k^{(0)}$ and δ. In each experiment, we fix two of the latter parameters, using the setup described in Section 4.3.2, and the matrix Φ, where its choices change with the experiments, and vary the remaining parameter. For each experiment, we ran 100 simulations, averaged the collusive levels among the 100 runs and computed the uncertainty levels using bootstrapping with a 99% confidence interval. Our figures present the averaged collusive level as a function of one parameter, where the shaded areas represent the uncertainty level.

Figure 4.8 investigates the dependence of the collusive level on the idiosyncratic preference parameter β_k, while considering two different choices of the externality matrix Φ: A symmetric one, where $\Phi = [1, 0; 0, 1]$ (left panel) and an asymmetric one, where $\Phi = [0, 1; -1, 0]$ (right panel). In both cases, we vary the idiosyncratic preference parameters and let $\beta_k \in [0.2, 6]$. In both panels, the collusive level sharply decreases when β_k is sufficiently small. In particular, the collusive level is high only when the degree of heterogeneity in users' tastes is sufficiently small. Section 4.5.1 interprets this behavior. We note that in the left panel, the sharp decrease stops when β_k exceeds 0.9, whereas in the right panel, this occurs once β_k exceeds 0.5. Furthermore, we note that in both subfigures, the collusive levels remain almost flat, at a value slightly below the collusive level $\Delta_0 = 0.3$, as the degree of heterogeneity in users' tastes exceeds 1. More experiments varying β_k with different choices of Φ are presented in Appendix Appendix C.3.

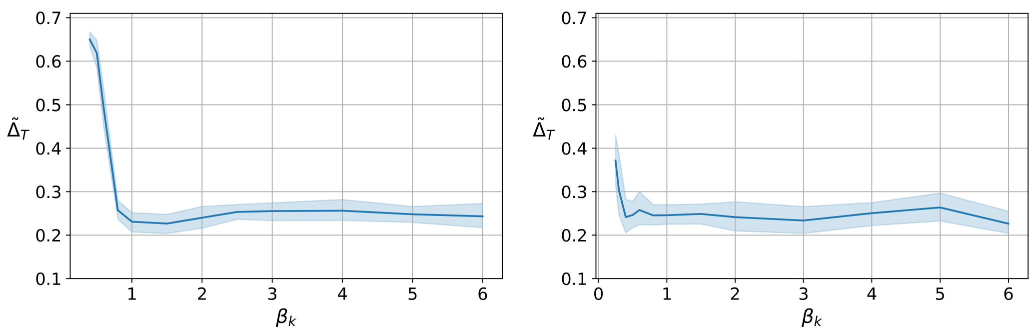


Figure 4.8: Collusive level with varying $\beta_b = \beta_s \in [0.2, 6]$ and different matrices Φ, where $\Phi = [1, 0; 0, 1]$ (left) and $\Phi = [0, 1; -1, 0]$ (right).

Figure 4.9 demonstrates the dependence of the collusive level on the outside option utility $u_k^{(0)}$ with the following choices for the externality matrix: $\Phi = [0, 0; 0, 0]$ (top left panel), $\Phi = [1, 0; 0, 1]$ (top right panel), $\Phi = [0, 1; 1, 0]$ (bottom left panel), and $\Phi = [1, 1; 1, 1]$ (bottom right panel). In all four panels, we observe a main trend of decrease of the collusive level as a function of the outside option utility over a sufficiently large domain. In all of these examples when $u_k^{(0)}$ is sufficiently small, the collusive level is at least 0.55, and when $u_k^{(0)}$ is sufficiently large, the collusive level is around 0.3, which is near the baseline level. The most significant reduction of the collusive level happens in a narrow range and the location of this significant decrease appears to be determined by the externalities as follows. It tends to move to the left when the network externalities are small and to the right when they are large. Additional examples in Figure C.4 support this conclusion. In addition, we observe that both left subfigures exhibit another small region of increase to the baseline level after the region of sharp decrease. This is not the case for the right subfigures. We also note a similar phenomenon in Figure C.4. It seems that an increase to the baseline level after a sharp decrease occurs in cases of sufficiently small externalities, where the threshold on externalities required to guarantee such a short increase is smaller for within-side externalities than for cross-side externalities. For example, considering cases of both Figures 4.9 and C.4, this short increase is observed at $\Phi = [0, 1; 1, 0]$ but not at $\Phi = [0, 2; 2, 0]$, and is observed at $\Phi = [1, 0; 0, 0]$ but not at $\Phi = [1, 0; 0, 1]$. Lastly, it follows from the bottom right subfigure of Figure 4.9 and cases of Figure C.4 that for sufficiently large externalities, the collusive level may slightly increase before the sharp decrease.

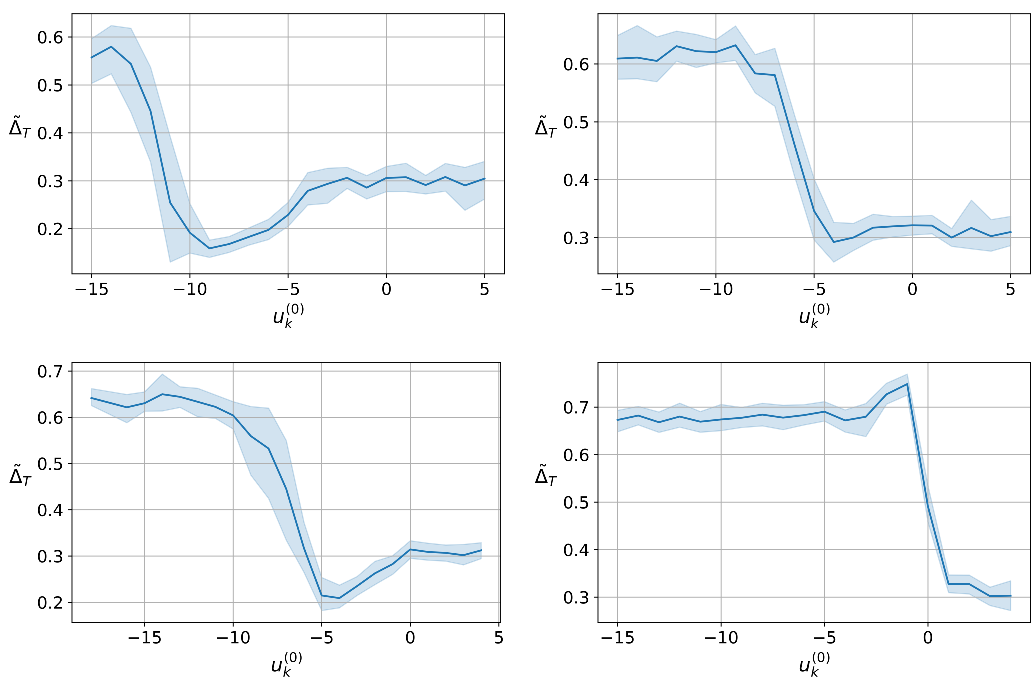


Figure 4.9: Collusive level with varying $u_k^{(0)}$, $k \in \{b, s\}$, with $\Phi = [0, 0; 0, 0]$ (top left), $\Phi = [1, 0; 0, 1]$ (top right), $\Phi = [0, 1; 1, 0]$ (bottom left) and $\Phi = [1, 1; 1, 1]$ (bottom right).

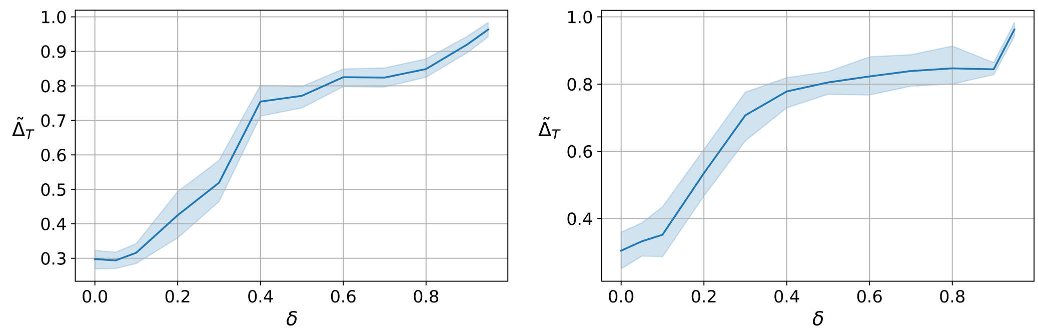


Figure 4.10: Collusive level with varying δ, where $\Phi = [0, 0; 0, 0]$ (left) and $\Phi = [1, 0; 0, 1]$ (right); in both cases $\delta \in [0.01, 0.99]$.

Figure 4.10 investigates the dependence of the collusive level on the discount rate δ, where $\delta \in [0.01, 0.99]$, with $\Phi = [0, 0; 0, 0]$ (left) and $\Phi = [1, 0; 0, 1]$ (right). In both panels, the collusive level increases with the discount rate. Thus, the more patient platforms are about the future, the higher the collusive level becomes. We note that the collusive level in the right panel is approximately a horizontal shift of that in the left panel. Thus, with a non-zero externality matrix the collusive level increases earlier than with the zero externality matrix. The positive relationship between the collusive level and the discount rate was first observed by Calvano et al. (2020b) for Bertrand model and our experiments verify the same relationship for a multi-sided market.

4.4.4 Discussion of Exceptional Cases

We numerically demonstrate two exceptional and uncommon scenarios: asymmetric optimal prices and competition prices larger than collusion prices.

Asymmetric collusion. Our metric for the collusive level compares the platform's rewards at time t with the symmetric equilibrium quantities π^* and π^C, following previous simulations of collusion (see, e.g., Calvano et al. (2020b) and Klein (2021)). However, our specific model may give rise to asymmetric equilibria and we thus study their possibility more carefully.

To allow asymmetric equilibria, we modify the definitions of the maximum values of the total profits and the collusive level as follows. The total profit, Π_a, is

$$\Pi_a(p_b^{(1)}, p_s^{(1)}, p_b^{(2)}, p_s^{(2)}) := \sum_{i=1}^{2}(x_b^{(i)}p_b^{(i)} + x_s^{(i)}p_s^{(i)}) \equiv \sum_{i=1}^{2}\pi_t^{(i)}, \tag{4.20}$$

where unlike (4.4) it does not assume symmetric prices (that is, it does not assume that $\pi_t^{(1)} = \pi_t^{(2)}$) and its subscript a indicates asymmetry. Similarly, the collusive level Δ_t is averaged among the two firms as follows

$$\Delta_t := (\Delta_t^{(1)} + \Delta_t^{(2)})/2 \equiv \frac{\Pi_a - 2\pi^*}{2(\pi^{\mathrm{C}} - \pi^*)},$$

where $\Delta_t^{(i)}$, $i \in \{1, 2\}$, was defined in (4.17).

Figure 4.11 demonstrates the maximum value that Δ_t can achieve in two controlled settings. In the first setting (left panel) $\Phi = [\phi_{bb}, \phi_{bs}; \phi_{bs}, \phi_{bb}]$ with $\phi_{bb} \in [-1, 1]$ and $\phi_{bs} \in [-4, 4]$, and in the second one (right panel) $\Phi = [\phi_{bb}, \phi_{bs}; -\phi_{bs}, \phi_{bb}]$ with $\phi_{bb} \in [-2, 2]$ and $\phi_{bs} \in [-8, 8]$. More precisely, we maximize Δ_t over $\boldsymbol{p}_t^{(1)} \in \mathcal{A}$ and $\boldsymbol{p}_t^{(2)} \in \mathcal{A}$, where $\mathcal{A}$ was defined in (4.13), and present the maximal values using a heatmap, whose values vary from purple (Δ_t greater than 1) to orange (Δ_t less than 1). In both panels, $\beta_k = 1.0$ and $u_k^{(0)} = -2.0$, $k \in \{b, s\}$.

In the left panel, the maximal Δ_t exceeds 1 when the within-side externality ϕ_{bb} is positive and close to 1 and the cross-side externality ϕ_{bs} is sufficiently negative. We note that it is much larger than 1. In the right panel, the maximal Δ_t is larger than 1 when ϕ_{bb} is sufficiently large and ϕ_{bs} is close to zero. We note though that it never exceeds the value of 1.0175. In both panels, for all other corresponding values of ϕ_{bb} and ϕ_{bs}, the maximum value of Δ_t is either achieved at a symmetric vector price or at an asymmetric vector price with corresponding maximum value close to 1. In the latter case, of the corresponding maximum value close to 1, our measure of collusive level $\tilde{\Delta}$ using the symmetric assumption can still be used to quantify the level of collusion. That is, one may often follow up our analysis with the symmetric quantities, except for special cases where the estimated collusive level exceeds one, where one needs to use the asymmetric quantities.

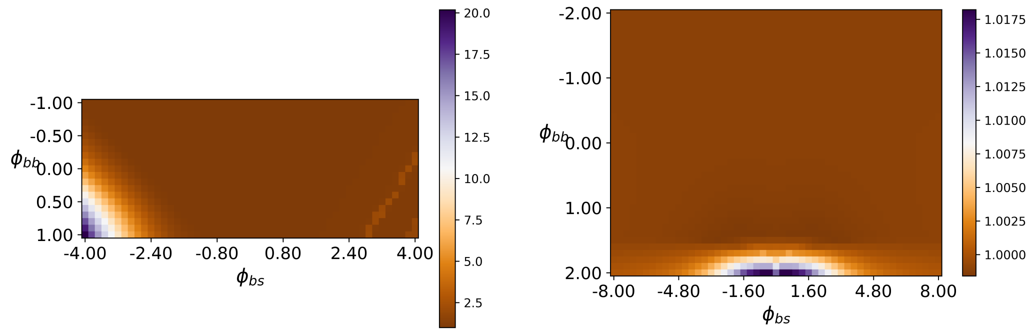


Figure 4.11: Maximal average collusive level with varying ϕ_{bb} and ϕ_{bs}, where $\Phi = [\phi_{bb}, \phi_{bs}; \phi_{bs}, \phi_{bb}]$ (left) and $\Phi = [\phi_{bb}, \phi_{bs}; -\phi_{bs}, \phi_{bb}]$ (right).

We remark that for almost all choices of β_k, $u_k^{(0)}$, δ and Φ in this paper, the asymmetric maximum value of Δ_t does not exceed the symmetric maximum value. Indeed, our simulations from Sections 4.4.1 through 4.4.3 show that $\tilde{\Delta}$ is smaller than or equal to 1. Nevertheless, when extending the right panel of Figure 4.6 to more negative values a collusive level higher than 1 is noticed, which we depict in Figure 4.12. In this figure, $\Phi = [0, \phi_{sb}; \phi_{sb}, 0]$ with $\phi_{sb} \in [-4, 2]$ and the asymmetric maximum value exceeds the symmetric maximum value when $\phi_{sb} < -2.5$. We notice that the collusive level reaches values close to 5 when the within-side externalities are zero and the cross-side externalities are equal in magnitude and sufficiently negative. Note that this scenario agrees with the one depicted in the left panel of Figure 4.11, where one can notice a similar value of 5 when $\phi_{bb} = \phi_{ss} = 0$ and $\phi_{bs} = \phi_{sb} = -4$.

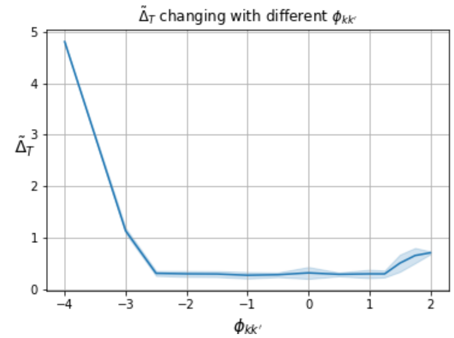


Figure 4.12: Collusive level with varying ϕ_{sb}, where $\Phi = [0, \phi_{sb}; \phi_{sb}, 0]$.

Competition prices larger than collusion prices. We recall that both the simulated action space presented in (4.12) and the formulated baseline model in Section 4.2.1 allow for the competition prices (p_k^*) to be either smaller or larger than the collusion prices (p_k^{C}). We demonstrate here an uncommon situation where the collusion price can be smaller than the competition price for one side of the market. In this example $\Phi = [1, -\phi_{sb}; \phi_{sb}, -2]$, $\beta_k = 0.5$ and $u_k^{(0)} = -1.0$, $k \in \{b, s\}$. If $\phi_{sb} \in [-5, 0.5)$, then $p_s^* < p_s^{\mathrm{C}}$. If $\phi_{sb} \in (0.5, 5]$, then $p_s^* > p_s^{\mathrm{C}}$. The left panel of Figure 4.13 demonstrates the competition and collusion prices of both sides of the market. For side b, the collusion price is always higher than the competition price. However, for side s, there are two different regimes separated by $\tilde{\phi}_{sb} \approx 0.5$. When $\phi_{sb} < \tilde{\phi}_{sb}$, the collusion price is higher than the competition price on the seller side, and when $\phi_{sb} > \tilde{\phi}_{sb}$ it is lower.

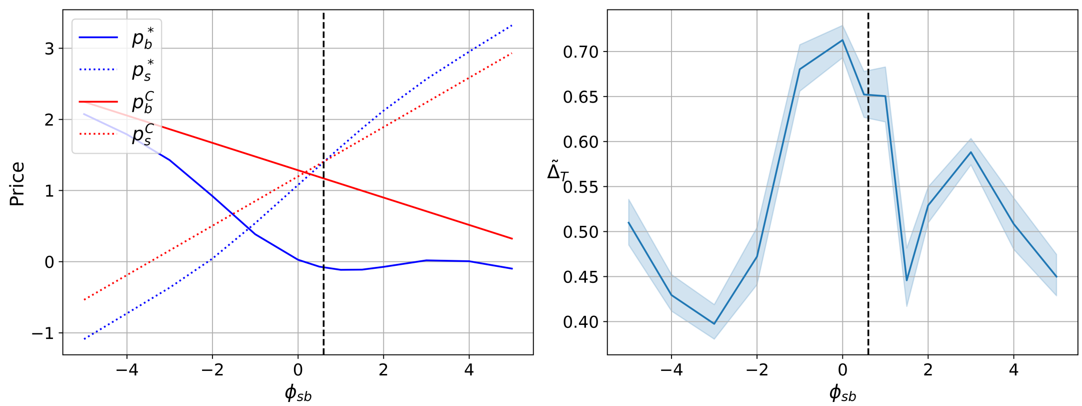


Figure 4.13: Demonstration of an uncommon scenario where the collusion prices can be lower than competition prices. Here, $\Phi = [1, -\phi_{sb}; \phi_{sb}, -2]$, where $\phi_{sb} \in [-5, 5]$. The left panel demonstrates collusion and competition prices on both sides of the market. The right panel demonstrates the collusive level. The black vertical dotted lines separate the two regimes: $p_s^* < p_s^{\mathrm{C}}$ on the left side of the black vertical dotted line and $p_s^* > p_s^{\mathrm{C}}$ on the right side of the black vertical dotted line.

The right panel of Figure 4.13 shows the collusive level for this example. The black dotted line is drawn at $\tilde{\phi}_{sb}$ to separate the two regimes. In the left regime, where $p_s^* < p_s^{\mathrm{C}}$, the collusive level decreases on $[-5, -3]$ and increases on $[-3, 0]$. This behavior is somewhat similar to the one described in the left panel of Figure 4.6. During the transition from $p_s^* < p_s^{\mathrm{C}}$ to $p_s^* > p_s^{\mathrm{C}}$, that is, in a small interval around $\tilde{\phi}_{sb}$, the collusive level decreases. Next, the collusive level increases on $[1.5, 3]$ and decreases on $[3, 5]$, which is the opposite behavior (in terms of decreasing and increasing) to the one demonstrated in the other regime.

4.5 Economic and Policy Discussion

We examine the economic implications of the numerical results presented in Section 4.4. Specifically, we analyze the impact of network externalities on the collusive level and discuss the ramifications of these findings on real-life markets. We then review the Preventing

Algorithmic Collusion Act of 2024, introduced by U.S. Senator Amy Klobuchar, and suggest an additional policy recommendation to ensure the safer use of Q-learning.

4.5.1 Economic Discussion

When $\Phi = 0$, our baseline platform competition game reduces to Bertrand competition games on each side of the market. For $\delta = 0.05$, $\beta_k = 1$ and $u_k^{(0)} = -2$, $k \in \{b, s\}$, our simulations show that Δ_0 in (4.19) is approximately 0.3. This means that when $\Phi = 0$, the profit gain relative to competition profits is about 30%. It is interesting to note that Calvano et al. (2020b) reported a different gain of approximately 20% for the same value of δ. To understand the difference, we note that in the latter paper, firms serve only one market and Bertrand competition is the baseline game. For that reason, the action space used in their simulations is smaller than the action space used in this simulation. We hypothesize that the higher-dimensional action space allows the platforms to mutually explore more information, increasing the chance of achieving a higher collusive level.

To test this hypothesis, we design the following experiment with a reduced action space dimension. We assume that the two sides of the markets are independent of each other (i.e., $\phi_{bs} = \phi_{sb} = 0$) and, assume that the platforms determine the price of a given side of the market based on the price history of this side only. That is, unlike our general setting, where $p_{t,k}^{(i)} = p_k^{(i)}(\boldsymbol{p}_{t-1,b}, \boldsymbol{p}_{t-1,s})$ for $k \in \{b, s\}$ (see (4.14)), we assume here that $p_{t,k}^{(i)} = p_k^{(i)}(\boldsymbol{p}_{t-1,k})$ for $k \in \{b, s\}$. The latter assumption reduces the dimension of the state space, and the goal of our experiment is to check whether this reduction in dimension decreases the collusive level.[9] Figure 4.14 investigates the dependence of the collusive level on the the discounting factor and the within side externalities in two controlled settings. In the first setting (left

[9]To be specific, in the previous numerical simulations in Section 4.4, the Q-function $Q(\boldsymbol{s}, \boldsymbol{a})$ is a function that maps $((\mathcal{P}_b \times \mathcal{P}_s) \times (\mathcal{P}_b \times \mathcal{P}_s)) \times (\mathcal{P}_b \times \mathcal{P}_s) \mapsto \mathbb{R}$, while the Q-function in this experiment has two independent component $Q_b(\boldsymbol{s}, \boldsymbol{a})$ and $Q_s(\boldsymbol{s}, \boldsymbol{a})$ that map $(\mathcal{P}_b \times \mathcal{P}_b) \times \mathcal{P}_b \mapsto \mathbb{R}$ and $(\mathcal{P}_s \times \mathcal{P}_s) \times \mathcal{P}_s \mapsto \mathbb{R}$ respectively. In other words, in the previous simulations, we needed to fit the tensor with $|\mathcal{P}_b|^3|\mathcal{P}_s|^3$ elements, whereas in the current simulation we fit two tensors whose number of elements are $|\mathcal{P}_b|^3$ and $|\mathcal{P}_s|^3$.

panel), $\Phi = [1, 0; 0, 1]$ and $\delta \in [0, 1]$, and in the second one (right panel), $\delta = 0.05$, $\Phi = [\phi_{bb}, 0; 0, \phi_{bb}]$ and $\phi_{bb} \in [-6, 2]$.

Comparing the left subfigure in Figure 4.14 and the right subfigure in Figure 4.10, we see that for each value of $\delta \in [0, 1]$, the collusive level is smaller in the former case. Similarly, comparing the right subfigure in Figure 4.14 and the right subfigure in Figure 4.5, we see that for each value of $\phi_{bb} \in [-6, 2]$, the collusive level is smaller in the former case. These two findings support our hypothesis: reducing the dimension of the action space when markets are independent, reduces the collusion level reached by the algorithm-driven platforms.

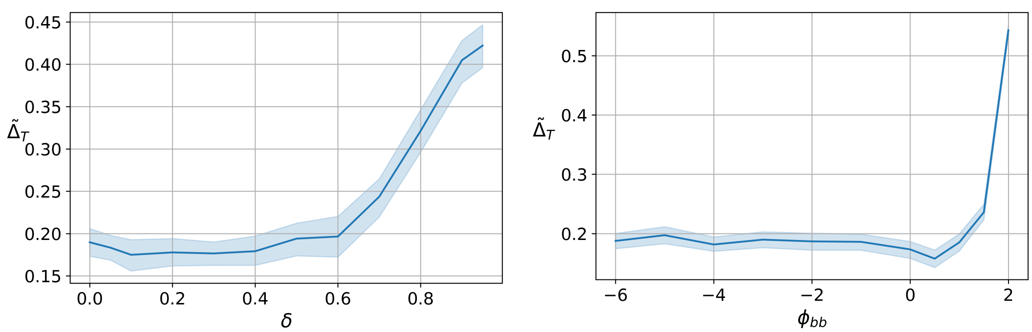


Figure 4.14: Collusive level with $\phi_{bs} = \phi_{sb} = 0$, $p^{(i)}_{t,k} = p^{(i)}_k(\boldsymbol{p}_{t-1,k})$ for $k \in \{b, s\}$. Left figure $\Phi = [1, 0; 0, 1]$ and varying $\delta \in [0, 1]$; right figure $\delta = 0.05$, $\Phi = [\phi_{bb}, 0; 0, \phi_{bb}]$ and varying $\phi_{bb} \in [-6, 2]$.

Next, we discuss the impact of $\Phi \neq 0$ on the collusive level compared to Bertrand competition, where $\Phi = 0$. When Φ has only one non-zero entry, the dependence of the collusive level on each possible entry is depicted in Figures 4.1 and the left panel of 4.5. We note that if the nonzero externality is either positive or sufficiently negative, then the collusion is higher in platform competition than in traditional Bertrand competition. Furthermore, when Φ is a diagonal matrix, the dependence of the collusive level on these entries is depicted in Figures 4.1 and the left panel of 4.2. We note that when these entries are both positive, collusion is higher in platform competition than in single-sided Bertrand competition. These findings suggest that in markets such as online or cloud gaming (e.g., *Xbox*, mobile and computer games), where positive within-side externalities are significant, algorithmic pricing will increase collusion levels beyond those observed in baseline Bertrand competition. Finally, when Φ has only off diagonal non-zero entries, the dependence on the collusive level on these entries is depicted in Figures 4.1 and the right panel of 4.2. We note that when these entries are both positive, collusion sharply increases above the baseline Bertrand level. These findings suggest that in markets such as video streaming (e.g., *Netflix*, *Hulu*, and *Amazon*) and social media markets (e.g., *Instagram* and *TikTok*), where positive cross-side externalities are significant, high levels of collusion can be expected if platforms use algorithmic pricing.

Traditionally, platforms exploit positive network externalities by subsidizing one market side to attract users on the other, boosting demand and profits (see, e.g., Armstrong and Wright (2007), Tan and Zhou (2021), and Chica et al. (2021)). Our findings suggest that algorithmic-driven platforms may also learn to leverage positive network externalities to significantly increase profit.

Next, we explore some cases where market participation remains constant while network externalities increase and examine whether the collusive level still increases and if so, by how much. For this purpose, we assume the default market parameters and two different settings of externalities: $\Phi = [\phi_{bb}, 0; 0, \phi_{bb}]$ with varying ϕ_{bb} (similar to the right panel of Figure 4.5) and $\Phi = [0, \phi_{bs}; \phi_{bs}, 0]$ with varying ϕ_{bs} (similar to the right panel of Figure 4.6). Furthermore, as we change Φ, we also adjust the value of the outside option utility $u_b^{(0)} = u_s^{(0)}$, so that the market participation remains constant and equal to the case where $\Phi = \mathbf{0}$ and $u_k^{(0)} = -2$ for $k \in \{b, s\}$.

Figure 4.15 demonstrates the collusive levels with varying Φ under the latter setting with the two different types of Φs. By comparing this Figure with the right panels of Figures 4.5 and 4.6, we recognize two important facts. First, the collusive level increases as either the within-side or cross-side externalities increase, even when market participation is constant. This means that the collusive level increases as the network effect increases and thus the high collusive level is generally not only attributed to mere increase of market participation. Second, the collusive level in Figure 4.15 is lower than the collusive levels achieved in the right panels of Figures 4.5 and 4.6. This indicates that the increasing market participation also contributes to the higher collusive level. However, the mere increase in network effect seems to be more impactful than the additional increase in market participation.

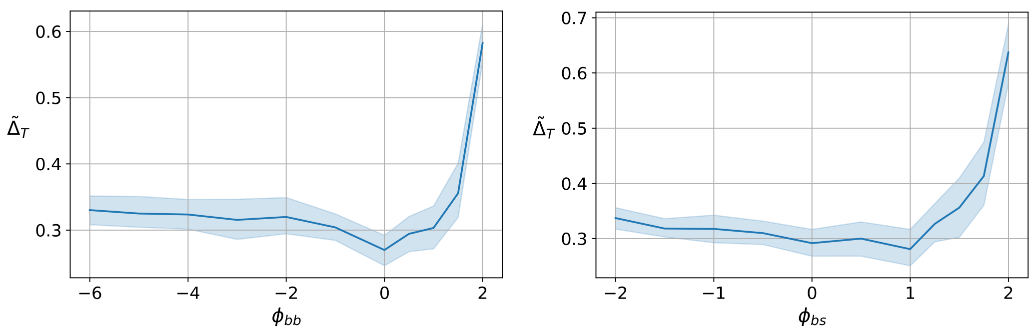


Figure 4.15: Collusive level when varying externalities, while keeping the market participation fixed (same as in the case where $\Phi = \mathbf{0}$ and $u_k^{(0)} = -2$ for $k = b, s$). Left: Collusive level with varying $\phi_{bb} = \phi_{ss}$ where $\Phi = [\phi_{bb}, 0; 0, \phi_{bb}]$; Right: Collusive level with varying $\phi_{bs} = \phi_{sb}$, where $\Phi = [0, \phi_{bs}; \phi_{bs}, 0]$.

Next we discuss the other scenario, where network externalities result in relatively small levels of collusion. First, we note that the right panels of both Figures 4.5 and 4.6 indicate examples where either the within-side externalities (Figure 4.5) or the cross-side externalities (Figure 4.6) are both negative with the same magnitude and in these cases the collusive level remains flat at the baseline competition level Δ_0. On the other hand, Figure 4.7 shows a case where the cross-side externality (ϕ_{bs}) is large and positive, and the within-side externality is sufficiently small and negative. In this scenario, the collusive level remains flat at a value slightly above Δ_0. The latter example is relevant to ride-sharing markets, where drivers compete with each other for riders, while riders benefit from faster pickup times. In this case, when using algorithmic pricing, our experiments indicate that collusive levels are close to the baseline level Δ_0.

Our findings reveal some interesting patterns in the dependence of the collusive level on three different market parameters: the degree of heterogeneity in users' tastes β_k, the constant term of the outside option utility $u_k^{(0)}$, and the discount rate δ. As shown by Figure 4.8, the collusive level sharply decreases as the degree of heterogeneity in users' tastes increases from 0.2 to 1. Afterwards, it remains flat around values lower than the baseline collusive level Δ_0. These findings can also be observed in Appendix Appendix C.3 for multiple choices of the externality matrix Φ. We are not aware of any previous result like this. A different result states that higher degree of heterogeneity in users' tastes leads to inelastic demand and higher price, which in turn leads to higher individual profits (see, e.g., Perloff and Salop (1985) and Anderson and De Palma (1992)). However, the collusion level generally does not correlate with individual profit values.

The main trend of the dependence of the collusive level on $u_k^{(0)}$ has relevant market implications. For instance, we note that for ride-sharing platforms using AI pricing, our analysis shows that an emphasis on increasing the outside option utility would decrease price and collusion levels. Such outside option utility can be increased by enhancing the public transportation system, and increasing mobility options such as e-bikes. Instead of investing further in such options and, in particular, in the safety of public transportation, the city of Minneapolis chose to decrease prices by passing an ordinance[10] that would force the two major ride-sharing platforms in the city, *Uber* and *Lyft*, to pay drivers the city's minimum hour wage. As a result, *Uber* and *Lyft* announced plans to leave the market. Our analysis indicates that there are other strategies to mitigate the problem.

Next, Figure 4.9 indicates a main trend of decreasing collusion levels as a function of the outside option utility. This behavior aligns with the observation of Chica et al. (2023b) (see Proposition 4.6) that as the value of the outside option increases, market power held

[10]The ordinance can be found at https://lims.minneapolismn.gov/Download/FileV2/32072/Transportation-Ride-Share-Worker-Protection-Ordinance.pdf

by platforms decreases. We further observed some localized trends, but they depend on specific choices of externalities. The main trend of collusion levels decreasing with $u_k^{(0)}$ has direct market implications. For instance, in the case of ride-sharing platforms using AI pricing, our analysis suggests that strengthening outside options–such as improving public transportation and expanding mobility alternatives like e-bikes–could reduce prices and collusion. Instead of prioritizing these investments, policymakers in Minneapolis initially sought to regulate ride-sharing prices through an ordinance mandating a minimum wage for drivers. This move prompted *Uber* and *Lyft* to announce their intent to exit the market, leading to a state-level compromise that increased driver wages while ensuring continued ride-sharing services. Our analysis suggests that while wage regulations can influence platform pricing dynamics, alternative strategies–such as improving transportation infrastructure–may provide a more sustainable approach to mitigating collusion, enhancing consumer welfare, and ensuring fairer fare distributions between drivers and platforms.

4.5.2 Policy Discussion

U.S. senator Amy Klobuchar introduced the *S.3686 - Preventing Algorithmic Collusion Act of 2024* in January 2024,[11] whose abstract is as follows:

> "*A bill to prevent anticompetitive conduct through the use of pricing algorithms by prohibiting the use of pricing algorithms that can facilitate collusion through the use of nonpublic competitor data, creating an antitrust law enforcement audit tool, increasing transparency, and enforcing violations through the Sherman Act and Federal Trade Commission Act, and for other purposes.*"

This bill reflects the increased concern by congress members and governmental institutions for the use of algorithmic price collusion. If such legislation succeeds, it will constitute a major advancement for consumer safety against the potential threats of AI. However,

[11] The act can be consulted at https://www.congress.gov/bill/118th-congress/senate-bill/3686/text,

the above act is rather limited. Indeed, it only targets AI algorithms trained with nonpublic competitor data and characterizes them as unlawful:

> " *SEC. 4. PREVENTING COLLUSIVE ACTIVITY IN PRICING ALGORITHMS. (a) In General.—It shall be unlawful for a person to use or distribute any pricing algorithm that uses, incorporates, or was trained with nonpublic competitor data.* "

Potentially collusive AI algorithms, such as the ones presented in this work using Q-learning, are left out of the scope of this act. We showed that these algorithms can learn to sustain high levels of collusion using only publicly available data, even in cases where agents have limited memory capacity. While this general observation has been established in previous works (Waltman and Kaymak (2008), Calvano et al. (2020b), Klein (2021) and Assad et al. (2024)), this work shows that in the presence of positive network externalities and two-sided markets, algorithmic driven platforms achieve collusive levels higher than those shown in previous works.

One way of decreasing collusion levels is obtained by increasing the value of the outside option utility, which has to be done in different ways for different markets (see e.g., the discussion in Section 4.5.1 for the ride-sharing platforms). Nevertheless, we next suggest a preliminary policy recommendation that can help avoid the risk of collusion by algorithmic driven AI agents in multiple markets. It is based on penalizing the reward in a reward-based method; for simplicity, we present it in the context of Q-learning.

Policy Recommendation (Q-learning with penalty term). We describe a very basic approach for reducing the collusive level by a potential intervention method. It can only be effective if it is enforced by regulators. The method introduces a penalty coefficient $\rho \geq 0$ that regularizes the Q-learning update formula as follows:

$$
\begin{aligned}
Q_{t+1}^{(i)}(\boldsymbol{s}_t^{(i)}, \boldsymbol{a}_t^{(i)}) &:= (1-\alpha) Q_t^{(i)}(\boldsymbol{s}_t^{(i)}, \boldsymbol{a}_t^{(i)}) \\
&\quad + \alpha \left(\pi_t^{(i)} + \delta \max_{\boldsymbol{a}} Q_t^{(i)}(\boldsymbol{s}_t^{(i)}, \boldsymbol{a}) - \rho \left((p_b^{(i)} - \bar{p}_b)_+ + (p_s^{(i)} - \bar{p}_s)_+ \right) \right),
\end{aligned}
$$

where $(p_k - \bar{p}_k)_+ = \max\{p_k - \bar{p}_k, 0\}$ and $\bar{p}_k = \frac{1}{N}\sum_{i=1}^{N} p_k^{(i)}$, for $k \in \{b, s\}$. Notice that the penalty term, $\rho\left((p_b^{(i)} - \bar{p}_b)_+ + (p_s^{(i)} - \bar{p}_s)_+\right)$, only becomes active, if either of the current prices is larger than the average price charged by all the platforms in the market at time t.

One should fix ρ to ensure a tolerable collusive level. For example, the strongest requirement of having no collusion fixes ρ such that $\tilde{\Delta}_T = 0$. If on the other hand, one accepts the collusion level with no externalities, i.e., when $\Phi = 0$, then one may fix ρ such that $\tilde{\Delta}_T = \Delta_0$. We note that any such chosen value of ρ is a function of the parameters of the model $\{\beta_k, \delta, \Phi, u_k^{(0)}\}$. Indeed, our findings suggest that any policy recommendation aimed to reduce the risk of algorithmic price collusion needs to be dependent on market parameters.

Figure 4.16 investigates the dependence of the collusive level on the penalty coefficient ρ for a setting with $\Phi = [2, 0; 0, 2]$ (left panel) and $\Phi = [0, 2; 2, 0]$ (right panel). Without penalty, these two cases have shown significantly high collusive levels (above 0.7) as shown in the right panel of Figure 4.5 and the right panel of Figure 4.6. In both cases, the collusive level reduces sharply as ρ increases and it reaches Δ_0 when ρ is approximately 0.2 (left panel) and 0.3 (right panel). For completeness, we also show extreme cases where the collusive level can be negative, which are different from our above recommendations for choosing ρ.

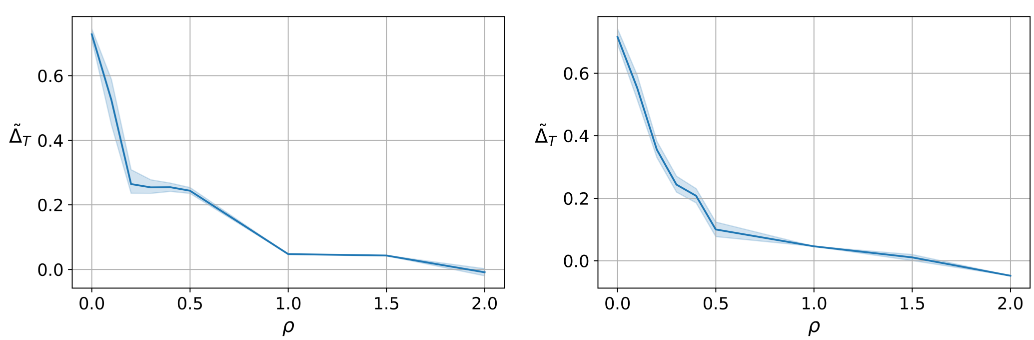


Figure 4.16: Collusive level with varying $\rho \in [0, 2]$: $\Phi = [2, 0; 0, 2]$ (left) and $\Phi = [0, 2; 2, 0]$ (right).

We note that this policy recommendation can be applied to any reward-based method by requiring platforms to include the penalty term $\rho\left((p_b(i) - \bar{p}_b)_+ + (p_s(i) - \bar{p}_s)_+\right)$ in their rewards. This recommendation is not limited to Q-learning.

This policy recommendation can be naturally extended to the scenario where platforms are not symmetric. We denote $c_b^{(i)}$ and $c_s^{(i)}$ as the costs for platform i on the buyer side and the seller side of the market, respectively. Furthermore, we denote $t_k^{(i)} := (p_k^{(i)} - c_k^{(i)})$, where $k \in \{b, s\}$, and rewrite the penalty for platform i as

$$\rho\left((t_b^{(i)} - \bar{t}_b)_+ + (t_s^{(i)} - \bar{t}_s)_+\right),$$

where $\bar{t}_k := \frac{1}{N}\sum_{i=1}^{N} t_k^{(i)}$.

4.6 Concluding Remarks

We developed a framework to explore algorithmic price collusion in platform competition and numerically analyzed how collusion levels depend on the externality matrix.

For zero network externalities, the profit gain relative to competitive profits is about 30%, exceeding the value reported by Calvano et al. (2020b) for the same discount factor δ. We attribute this to the higher-dimensional action space, which enables platforms to mutually explore more information, increasing the likelihood of achieving high collusion levels.

In common economic scenarios with positive network externalities–particularly when these externalities appear in single entries, diagonal elements, or off-diagonal elements of Φ–collusion in platform competition is significantly stronger than in traditional Bertrand competition. Our analysis suggests that this high collusive level stems primarily from strong network interaction effects, with a secondary contribution from increased market participation. This implies that in markets such as online gaming, video streaming, and social media, AI-driven platforms may leverage positive network externalities to enhance profits through algorithmic pricing, raising concerns about high levels of collusion.

Additionally, our findings reveal patterns in how collusion depends on user heterogeneity, discount rates, and outside option utility. Specifically, greater heterogeneity in user preferences reduces collusion. Higher discount rates lead to increased collusion, consistent with Calvano et al. (2020b) for Bertrand competition, though the rate of increase is steeper in platform settings. We provide evidence in the appendix that this collusion remains tacit. Moreover, higher outside option utility generally suppresses collusion within a sufficiently large domain. This suggests that regulators could mitigate collusion by enhancing outside options–such as improving public transportation in ride-sharing markets.

To address algorithmic price collusion, we also propose a variant of Q-learning with a penalty term. It can be used to inform policy interventions based on market parameters.

Several future research directions remain open. First, while Chica et al. (2023b) models platform competition with an unlimited number of platforms, our numerical experiments were limited to two platforms due to the exponential growth of the state space. Future work could explore alternative Q-function estimation methods to manage these computational challenges. Second, it would be valuable to extend our experiments to models incorporating nonlinear network externality effects and multi-homing, where users participate in multiple platforms.

References

Anderson, S. P. and De Palma, A. (1992). The logit as a model of product differentiation. *Oxford Economic Papers*, 44(1):51–67. 39, 49, 58, 64, 104

Anderson, S. P., De Palma, A., and Thisse, J.-F. (1992). *Discrete choice theory of product differentiation*. MIT press. 44

Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica: Journal of the Econometric Society*, pages 51–75. 47

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691. 33, 34, 44, 72

Armstrong, M. and Vickers, J. (2001). Competitive price discrimination. *The RAND Journal of Economics*, pages 579–605. 31, 46

Armstrong, M. and Wright, J. (2007). Two-sided markets, competitive bottlenecks and exclusive contracts. *Economic Theory*, 32:353–380. 102

Arslantas, Y., Yuceel, E., and Sayin, M. O. (2024). Strategizing against q-learners: A control-theoretical approach. *arXiv preprint arXiv:2403.08906*. 72

Assad, S., Clark, R., Ershov, D., and Xu, L. (2024). Algorithmic pricing and competition: Empirical evidence from the German retail gasoline market. *Journal of Political Economy*, 132(3):000–000. 5, 69, 71, 106

Bardey, D., Cremer, H., and Lozachmeur, J.-M. (2014). Competition in two-sided markets with common network externalities. *Review of Industrial Organization*, 44:327–345. 47

Barlo, M., Carmona, G., and Sabourian, H. (2009). Repeated games with one-memory. *Journal of Economic Theory*, 144(1):312–336. 6, 8, 70

Barlo, M., Carmona, G., and Sabourian, H. (2016). Bounded memory folk theorem. *Journal of economic theory*, 163:728–774. 8, 197

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262. 39

Bertomeu, J., Evans III, J. H., Feng, M., and Tseng, A. (2021). Tacit collusion and voluntary disclosure: Theory and evidence from the US automotive industry. *Management science*, 67(3):1851–1875. 69

Besanko, D., Gupta, S., and Jain, D. (1998). Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Science*, 44(11-part-1):1533–1547. 39

Bhatia, R. (2013). *Matrix analysis*, volume 169. Springer Science & Business Media. 133

Bishop, R. L. (1960). Duopoly: Collusion or warfare? *The American Economic Review*, 50(5):933–961. 56

Brander, J. A. and Spencer, B. J. (1985). Tacit collusion, free entry and welfare. *The Journal of Industrial Economics*, pages 277–294. 56

Brero, G., Mibuari, E., Lepore, N., and Parkes, D. C. (2022). Learning to mitigate ai collusion on economic platforms. *Advances in Neural Information Processing Systems*, 35:37892–37904. 71

Brown, Z. Y. and MacKay, A. (2023). Competition in pricing algorithms. *American Economic Journal: Microeconomics*, 15(2):109–156. 5, 72

Caillaud, B. and Jullien, B. (2003). Chicken & egg: Competition among intermediation service providers. *The RAND journal of Economics*, pages 309–328. 30, 32, 33

Calvano, E., Calzolari, G., Denicolò, V., Harrington Jr, J. E., and Pastorello, S. (2020a). Protecting consumers from collusive prices due to ai. *Science*, 370(6520):1040–1042. 69

Calvano, E., Calzolari, G., Denicolo, V., and Pastorello, S. (2020b). Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, 110(10):3267–97. 1, 3, 5, 22, 24, 26, 69, 70, 71, 81, 95, 100, 106, 108, 109, 199

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 785–794. 83

Chica, C., Chuk, K., and Tamayo, J. A. (2021). *Exclusive Dealing and Entry by Competing Two-Sided Platforms*. Harvard Business School. 34, 42, 44, 47, 50, 68, 72, 102

Chica, C., Guo, Y., and Lerman, G. (2023a). Competition and collusion in two-sided markets with an outside option. *Available at SSRN 4641282*. 16, 25, 27

Chica, C., Guo, Y., and Lerman, G. (2023b). Competition and collusion in two-sided markets with an outside option. *Available at SSRN 4641282*. 72, 74, 75, 79, 104, 109

Chica, C., Guo, Y., and Lerman, G. (2024a). Artificial intelligence and algorithmic price collusion in two-sided markets. *arXiv preprint arXiv:2407.04088*. 5, 22, 24, 26, 66, 68

Chica, C., Guo, Y., and Lerman, G. (2024b). Stochastic games, reinforcement learning and collusion with bounded memory. In preparation. 197

Cohen, M. C. and Zhang, R. (2022). Competition and coopetition for two-sided platforms. *Production and Operations Management*, 31(5):1997–2014. 34, 57

Conlon, C. and Gortmaker, J. (2020). Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics*, 51(4):1108–1161. 39

Correia-da Silva, J., Jullien, B., Lefouili, Y., and Pinho, J. (2019). Horizontal mergers between multisided platforms: Insights from cournot competition. *Journal of Economics & Management Strategy*, 28(1):109–124. 35

Dewenter, R., Haucap, J., and Wenzel, T. (2011). Semi-collusion in media markets. *International Review of Law and Economics*, 31(2):92–98. 16, 25, 35, 57

Du, K. and Tanriverdi, H. (2023). Does it enable collusion or competition: Examining the effects of it on service pricing in multimarket multihospital systems. *MIS Quarterly*, 47(4). 69

Fink, A. M. (1964). Equilibrium in a stochastic n-person game. *Journal of science of the hiroshima university, series ai (mathematics)*, 28(1):89–93. 4, 5, 6, 12, 13, 14, 15, 29, 132, 134

Friedman, J. W. (1985). Cooperative equilibria in finite horizon noncooperative supergames. *Journal of Economic Theory*, 35(2):390–398. 17, 18

Gal-Or, E. (2020). Market segmentation on dating platforms. *International Journal of Industrial Organization*, 68:102558. 31

Gilbert, E. M. (2019). Antitrust and commitment issues: Monopolization of the dating app industry. *NYUL Rev.*, 94:862. 66

Halaburda, H., Jan Piskorski, M., and Yıldırım, P. (2018). Competing by restricting choice: The case of matching platforms. *Management Science*, 64(8):3574–3594. 31

Hsu, J. and Wang, X. H. (2005). On welfare under cournot and bertrand competition in differentiated oligopolies. *Review of Industrial Organization*, 27:185–191. 61

Hu, J. and Wellman, M. P. (2003). Nash q-learning for general-sum stochastic games. *Journal of machine learning research*, 4(Nov):1039–1069. 16

Jaakkola, T., Jordan, M., and Singh, S. (1993). Convergence of stochastic iterative dynamic programming algorithms. *Advances in neural information processing systems*, 6. 5

Jeitschko, T. D. and Tremblay, M. J. (2020). Platform competition with endogenous homing. *International Economic Review*, 61(3):1281–1305. 35

Johnson, J. P., Rhodes, A., and Wildenbeest, M. (2023). Platform design when sellers use pricing algorithms. *Econometrica*, 91(5):1841–1879. 71

Jullien, B. and Pavan, A. (2019). Information management and pricing in platform markets. *The Review of Economic Studies*, 86(4):1666–1703. 34

Klein, T. (2021). Autonomous algorithmic collusion: Q-learning under sequential pricing. *The RAND Journal of Economics*, 52(3):538–558. 5, 22, 26, 69, 71, 81, 95, 106

Lefouili, Y. and Pinho, J. (2020). Collusion between two-sided platforms. *International Journal of Industrial Organization*, 72:102656. 35

Maskin, E. and Tirole, J. (1988). A theory of dynamic oligopoly, ii: Price competition, kinked demand curves, and edgeworth cycles. *Econometrica: Journal of the Econometric Society*, pages 571–599. 201

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49. 15, 119

Obara, I. and Zincenko, F. (2017). Collusion and heterogeneity of firms. *The RAND Journal of Economics*, 48(1):230–249. 71

OECD (2017). Algorithms and collusion: Competition policy in the digital age. *Technical Report*. 5, 69

Ok, E. A. (2007). *Real analysis with economic applications*, volume 10. Princeton University Press. 135

Osborne, M. J. (1994). *A course in game theory*. MIT Press. 18

Ozdaglar, A., Sayin, M. O., and Zhang, K. (2021). Independent learning in stochastic games. *arXiv preprint arXiv:2111.11743*. 6

Peitz, M. and Sato, S. (2023). Asymmetric platform oligopoly. Technical report, University of Bonn and University of Mannheim, Germany. 35

Perloff, J. M. and Salop, S. C. (1985). Equilibrium with product differentiation. *The Review of Economic Studies*, 52(1):107–120. 47, 104

Possnig, C. (2023). *Reinforcement learning and collusion*. Department of Economics, University of Waterloo. 6

Rochet, J.-C. and Choné, P. (1998). Ironing, sweeping, and multidimensional screening. *Econometrica*, pages 783–826. 47

Rochet, J.-C. and Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029. 30, 33

Rochet, J.-C. and Tirole, J. (2006). Two-sided markets: a progress report. *The RAND Journal of Economics*, 37(3):645–667. 33

Ruhmer, I. (2010). Platform collusion in two-sided markets. 72, 198

Ryan, C. (2017). Computer and internet use in the United States: 2016. *American Community Survey Reports, ACS-39, U.S. Census Bureau, Washington, DC*. 30

SSRS (2024). The public and online dating in 2024. 31

Tan, G. and Zhou, J. (2021). The effects of competition and entry in multi-sided markets. *The Review of Economic Studies*, 88(2):1002–1030. 31, 33, 34, 36, 42, 44, 47, 54, 59, 61, 63, 72, 102

Tan, H. and Wright, J. (2021). Pricing distortions in multi-sided platforms. *International Journal of Industrial Organization*, 79:102732. 49

Teh, T.-H., Liu, C., Wright, J., and Zhou, J. (2023). Multihoming and oligopolistic platform competition. *American Economic Journal: Microeconomics*, 15(4):68–113. 35, 68

Tirole, J. (1988). *The theory of industrial organization*. MIT Press. 16, 25, 64, 71

Tremblay, M. J., Adachi, T., and Sato, S. (2023). Cournot platform competition with mixed-homing. *International Journal of Industrial Organization*, 91:103002. 35

Varian, H. R. (1989). Price discrimination. *Handbook of industrial organization*, 1:597–654. 47, 56

Waltman, L. and Kaymak, U. (2008). Q-learning agents in a cournot oligopoly model. *Journal of Economic Dynamics and Control*, 32(10):3275–3293. 1, 5, 71, 106

Watkins, C. J. and Dayan, P. (1992). Q-learning. *Machine learning*, 8(3):279–292. 5, 18, 23, 80

Wells, J. R., Danskin, G., and Ellsworth, G. (2015; revised in 2021). Amazon.com, 2021. *Harvard Business School Case Study 716-402*. 30

Werden, G. J., Froeb, L. M., and Tardiff, T. J. (1996). The use of the logit model in applied industrial organization. *International Journal of the Economics of Business*, 3(1):83–105. 39

Weyl, E. G. (2010). A price theory of multi-sided platforms. *American Economic Review*, 100(4):1642–1672. 33

White, A. and Weyl, E. G. (2016). Insulated platform competition. *Available at SSRN 1694317*. 33, 34, 36, 72

Appendix A

Proofs of Chapter 2

A.1 Proof of Theorem 2

We start by proving that for each $(i, s_1, \boldsymbol{p}_0)$-coordinate

$$\underbrace{\max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \sigma_1^i, \boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0}}_{LHS} = \underbrace{\max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i | \boldsymbol{\sigma}_1^{-i*})}_{RHS}.$$

We first prove that LHS $\leq$ RHS and then that LHS $\geq$ RHS.

Proof of LHS $\leq$ RHS: Since $\boldsymbol{v}^*$ satisfies equations (2.16) and (2.17), for each $(i, s_1, \boldsymbol{p}_0)$-coordinate

$$\max_{\sigma_1^i} \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \sigma_1^i, \boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0} = \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \boldsymbol{\sigma}_1^*, \boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0}. \tag{A.1}$$

From (2.10), (A.6) and Proposition 1,

$$\begin{aligned} &\boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \boldsymbol{\sigma}_1^*, \boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0} \\ &= \sum_{\boldsymbol{p}_1 \in \mathcal{A}^n} \sigma_1^*(\boldsymbol{p}_1 | \boldsymbol{p}_0, s_1) \left[\pi^i(\boldsymbol{p}_1, s_1) + \delta_i \sum_{s_2 \in \mathcal{S}} \mathbb{P}(s_2 | \boldsymbol{p}_1, s_1) \tilde{V}_1^i(s_2, \boldsymbol{p}_1, \sigma_1^{i*} | \boldsymbol{\sigma}_1^{-i*}) \right] \\ &= \tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^{i*} | \boldsymbol{\sigma}_1^{-i*}). \end{aligned} \tag{A.2}$$

Clearly, (A.1) and (A.2) imply that LHS$\leq$RHS.

Proof of LHS $\geq$ RHS: For each coordinate $(i, s_1, \boldsymbol{p}_0)$ and $\sigma_1^i \in \boldsymbol{\Sigma}_1^i$, we estimate the following quantity,

$$\tilde{V}_1^i(s_1,\boldsymbol{p}_0,\sigma_1^i|\boldsymbol{\sigma}_1^{-i*}) - \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*,\sigma_1^i,\boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0} = \sum_{\boldsymbol{p}_1\in\mathcal{A}^n}\sigma_1^i(p_1^i|\boldsymbol{p}_0,s_1)\sigma_1^{-i*}(\boldsymbol{p}_1^{-i}|\boldsymbol{p}_0,s_1)$$
$$\cdot\,\delta_i\sum_{s_2\in\mathcal{S}}\mathbb{P}(s_2|\boldsymbol{p}_1,s_1)(\tilde{V}_1^i(s_2,\boldsymbol{p}_1,\sigma_1^i|\boldsymbol{\sigma}_1^{-i*}) - \tilde{V}_1^i(s_2,\boldsymbol{p}_1,\sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*})). \tag{A.3}$$

We have used equation (2.16), which claims that $\boldsymbol{v}^* = \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*,\boldsymbol{\sigma}_1^*,\boldsymbol{v}^*)$ and we have used equation (A.2). We denote

$$\Delta\boldsymbol{V}_1^i(s_2,\boldsymbol{p}_1,\sigma_1^i,\boldsymbol{\sigma}_1^*) := \tilde{V}_1^i(s_2,\boldsymbol{p}_1,\sigma_1^i|\boldsymbol{\sigma}_1^{-i*}) - \tilde{V}_1^i(s_2,\boldsymbol{p}_1,\sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*})$$

. Applying first the fact that $-\tilde{V}_1^i(s_1,\boldsymbol{p}_0,\sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*}) \le -\boldsymbol{V}_1(\boldsymbol{\sigma}_1^*,\sigma_1^i,\boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0}$ (which follows from (A.1) and (A.2)) and then (A.3) result in

$$\begin{aligned}&\Delta\boldsymbol{V}_1^i(s_1,\boldsymbol{p}_0,\sigma_1^i,\boldsymbol{\sigma}_1^*)\\ &\le \textstyle\sum_{\boldsymbol{p}_1\in\mathcal{A}^n}\sigma_1^i(p_1^i|\boldsymbol{p}_0,s_1)\sigma_1^{-i*}(\boldsymbol{p}_1^{-i}|\boldsymbol{p}_0,s_1)\delta_i\sum_{s_2\in\mathcal{S}}\mathbb{P}(s_2|\boldsymbol{p}_1,s_1)\max_{(s_2,\boldsymbol{p}_1)\in\mathcal{S}\times\mathcal{A}^n}\Delta\boldsymbol{V}_1^i(s_2,\boldsymbol{p}_1,\sigma_1^i,\boldsymbol{\sigma}_1^*)\\ &= \delta_i\max_{(s_2,\boldsymbol{p}_1)\in\mathcal{S}\times\mathcal{A}^n}\Delta\boldsymbol{V}_1^i(s_2,\boldsymbol{p}_1,\sigma_1^i,\boldsymbol{\sigma}_1^*).\end{aligned} \tag{A.4}$$

Since (A.4) holds for all $(s_1,\boldsymbol{p}_0)\in\mathcal{S}\times\mathcal{A}^n$ and $\delta_i<1$

$$\max_{(s_1,\boldsymbol{p}_0)\in\mathcal{S}\times\mathcal{A}^n}\Delta\boldsymbol{V}_1^i(s_1,\boldsymbol{p}_0,\sigma_1^i,\boldsymbol{\sigma}_1^*)\le 0.$$

That is, $\tilde{V}_1^i(s_1,\boldsymbol{p}_0,\sigma_1^i|\boldsymbol{\sigma}_1^{-i*}) \le \tilde{V}_1^i(s_1,\boldsymbol{p}_0,\sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*})$ for each $(s_1,\boldsymbol{p}_0)\in\mathcal{S}\times\mathcal{A}^n$ and $\sigma_1^i\in\boldsymbol{\Sigma}_1^i$. We thus conclude that LHS $\ge$ RHS.

Lastly, we show that $\boldsymbol{\sigma}_1^*$ is a Nash equilibrium from time $t=1$. Fix $i\in[n]$. By (2.10), equation (2.16) yields for each $(s_1,\boldsymbol{p}_0)\in\mathcal{S}\times\mathcal{A}^n$,

$$v^*_{i,s_1,\boldsymbol{p}_0} = \sum_{\boldsymbol{p}_1\in\mathcal{A}^n}\sigma_1^*(\boldsymbol{p}_1|\boldsymbol{p}_0,s_1)\left[\pi^i(\boldsymbol{p}_1,s_1)+\delta_i\sum_{s_2\in\mathcal{S}}\mathbb{P}(s_2|\boldsymbol{p}_1,s_1)v^*_{i,s_2,\boldsymbol{p}_1}\right]. \tag{A.5}$$

By Proposition 1, the sequence $\{\tilde{V}_1^i(s_1,\boldsymbol{p}_0,\sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*})\}_{(s_1,\boldsymbol{p}_0)\in\mathcal{S}\times\mathcal{A}^n}$ is the unique solution to the system described by (A.5). Therefore, for each $(s_1,\boldsymbol{p}_0)\mathcal{S}\times\mathcal{A}^n$

$$v^*_{i,s_1,\boldsymbol{p}_0} = \tilde{V}_1^i(s_1,\boldsymbol{p}_0,\sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*}). \tag{A.6}$$

By (2.17) and (A.6),

$$\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*}) = \max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \sigma_1^i, \boldsymbol{v}^*)_{i,s_1,\boldsymbol{p}_0}. \tag{A.7}$$

By (2.18), which we proved above,

$$\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*}) = \max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i|\boldsymbol{\sigma}_1^{-i*}). \tag{A.8}$$

It follows that $\boldsymbol{\sigma}_1^*$ is a Nash equilibrium from $t = 1$. □

A.2 Proof Theorem 3

Let $\boldsymbol{\sigma}_1^* \in \boldsymbol{\Sigma}_1$ and $\boldsymbol{v}^* \in \mathbb{R}^{nrM}$ be the quantities given by Theorem 1. By Theorem 2, $\boldsymbol{\sigma}_1^*$ is a Nash equilibrium from time $t = 1$. To prove the theorem, we need to show that there exists $\boldsymbol{\sigma}_0^* \in \boldsymbol{\Sigma}_0$ satisfying for each $i \in [n]$

$$\sigma_0^{i*} \in \text{argmax}_{\sigma_0^i \in \boldsymbol{\Sigma}_0} \tilde{V}_0^i(s_0, (\sigma_0^i, \sigma_1^{i*})|(\boldsymbol{\sigma}_0^{-i*}, \boldsymbol{\sigma}_1^{-i*})). \tag{A.9}$$

We can rewrite the above equation by defining for each $(\boldsymbol{p}_0, s_0) \in \mathcal{A}^n \times \mathcal{S}$

$$Q^{i*}(\boldsymbol{p}_0, s_0) := \pi^i(\boldsymbol{p}_0, s_0) + \delta_i \sum_{s_1 \in \mathcal{S}} \mathbb{P}(s_1|\boldsymbol{p}_0, s_0) v_{i,s_1,\boldsymbol{p}_0}^* \tag{A.10}$$

and noting that

$$\tilde{V}_0^i(s_0, (\sigma_0^i, \sigma_1^{i*})|(\boldsymbol{\sigma}_0^{-i*}, \boldsymbol{\sigma}_1^{-i*})) = \mathbb{E}_{(\sigma_0^i, \boldsymbol{\sigma}_0^{-i*})}[Q^{i*}(\boldsymbol{p}, s)|s_0]. \tag{A.11}$$

By Theorem 1 and equation (2.12), $\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*}) = v^*_{i,s_1,\boldsymbol{p}_0}$. Using the latter fact, and (2.6), (2.8) and (2.9) we prove (A.11) by obtaining for each $s_0 \in \mathcal{S}$ and $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_1^*)$

$$\begin{aligned}\tilde{V}_0^i(s_0, \boldsymbol{\sigma}^i|\boldsymbol{\sigma}^{-i}) &= \sum_{\boldsymbol{p}_0\in\mathcal{A}^n} \sigma_0(\boldsymbol{p}_0|s_0)\left\{\pi^i(\boldsymbol{p}_0, s_0) + \delta_i \sum_{s_1\in\mathcal{S}} \mathbb{P}(s_1|\boldsymbol{p}_0, s_0)\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^{i*}|\boldsymbol{\sigma}_1^{-i*})\right\}\\ &= \sum_{\boldsymbol{p}_0\in\mathcal{A}^n} \sigma_0(\boldsymbol{p}_0|s_0)\left\{\pi^i(\boldsymbol{p}_0, s_0) + \delta_i \sum_{s_1\in\mathcal{S}} \mathbb{P}(s_1|\boldsymbol{p}_0, s_0)v^*_{i,s_1,\boldsymbol{p}_0}\right\}\\ &= \mathbb{E}_{\boldsymbol{\sigma}_0}[Q^{i*}(\boldsymbol{p}, s)|s_0].\end{aligned} \tag{A.12}$$

The use of (A.11) in (A.9) easily concludes the proof. Indeed, the existence of $\boldsymbol{\sigma}_0^* \in \boldsymbol{\Sigma}_0$ satisfying for each $i \in [n]$

$$\sigma_0^{i*} \in \text{argmax}_{\sigma_0^i\in\boldsymbol{\Sigma}_0}\mathbb{E}_{(\sigma_0^i,\boldsymbol{\sigma}_0^{-i*})}[Q^{i*}(\boldsymbol{p}, s)|s_0].$$

is guaranteed by the existence of Nash equilibrium in mixed strategies in Nash (1950). The profile $(\boldsymbol{\sigma}_0^*, \boldsymbol{\sigma}_1^*)$, where $\boldsymbol{\sigma}_1^*$ is given by Theorem 2 and $\boldsymbol{\sigma}_0^*$ is given by (2.15), is a one-memory SPE of the stochastic game. □

A.3 Proof of Proposition 2

Recall that each firm uses $\boldsymbol{\sigma}^f = (\sigma_0^f, \sigma_1^f)$, where $\sigma_0^f(p^C) = 1$, $\sigma_1^f(p^C|\boldsymbol{p}^C) = 1$, and $\sigma_1^f(p^*|\boldsymbol{p}_0) = 1$ for each $\boldsymbol{p}_0 \in \mathcal{A}^n \setminus \{\boldsymbol{p}^C\}$. We use Algorithm 1 to show that $\boldsymbol{\sigma}^f$ is an SPE of the stochastic game.

Step 1 of Algorithm 1: We plug $\boldsymbol{\sigma}_1^f$ into equation (2.16) and solve it as a linear system with unknowns listed in the vector $\boldsymbol{v}^f = (v^f_{i,\boldsymbol{p}_0})_{i\in[n],\boldsymbol{p}_0\in\mathcal{A}^n}$, and obtain

$$v^f_{i,\boldsymbol{p}_0} = \boldsymbol{V}_1(\boldsymbol{\sigma}_1^f, \boldsymbol{\sigma}_1^f, \boldsymbol{v}^f)_{i,\boldsymbol{p}_0}. \tag{A.13}$$

By (2.10), (A.13) is equivalent to

$$v^f_{i,\boldsymbol{p}_0} = \sum_{\boldsymbol{p}_1\in\mathcal{A}^n} \sigma_1^f(\boldsymbol{p}_1|\boldsymbol{p}_0)\left[\pi^i(\boldsymbol{p}_1) + \delta_i v^f_{i,\boldsymbol{p}_1}\right].$$

It follows that for each $i \in [n]$,

$$v^f_{i,\boldsymbol{p}_0} = \frac{1}{1-\delta_i} \cdot \begin{cases} \pi^i(\boldsymbol{p}^C) & \text{if } \boldsymbol{p}_0 = \boldsymbol{p}^C, \\ \pi^i(\boldsymbol{p}^*) & \text{if } \boldsymbol{p}_0 \neq \boldsymbol{p}^C. \end{cases} \tag{A.14}$$

Step 2 of Algorithm 1: We plug $\boldsymbol{v}^f$ and $\boldsymbol{\sigma}^f_1$ into (2.17) and show that $\boldsymbol{v}^f$ is a fixed point of the operator $v_{i,\boldsymbol{p}_0} \mapsto \max_{\sigma^i_1 \in \boldsymbol{\Sigma}^i_1} \boldsymbol{V}_1(\boldsymbol{\sigma}^f_1, \sigma^i_1, \boldsymbol{v})_{i,\boldsymbol{p}_0}$. By Assumption 2, $\boldsymbol{p}^*$ is a Nash equilibrium of the game $(\pi^i(\cdot))^n_{i=1}$, and thus

$$\frac{\pi^i(\boldsymbol{p}^*)}{1-\delta_i} \geq \max_{p^i \in \mathcal{A}\setminus\{p^*\}} \pi^i(p^i, (\boldsymbol{p}^*)^{-i}) + \delta_i \frac{\pi^i(\boldsymbol{p}^*)}{1-\delta_i}. \tag{A.15}$$

Similarly, by rewriting Assumption 3-(ii), we obtain

$$\frac{\pi^i(\boldsymbol{p}^C)}{1-\delta_i} \geq \max_{p^i \in \mathcal{A}\setminus\{p^C\}} \pi^i(p^i, (\boldsymbol{p}^C)^{-i}) + \delta_i \frac{\pi^i(\boldsymbol{p}^*)}{1-\delta_i}. \tag{A.16}$$

By (A.14), (A.15) and (A.16), it follows that

$$\begin{aligned}
&\max_{\tau^i_1 \in \boldsymbol{\Sigma}^i_1} \boldsymbol{V}_1(\boldsymbol{\sigma}^f_1, \tau^i_1, \boldsymbol{v}^f)_{i,\boldsymbol{p}_0} \\
&= \max_{\tau^i_1 \in \boldsymbol{\Sigma}^i_1} \sum_{p^i \in \mathcal{A}} \tau^i_1(p^i|\boldsymbol{p}_0) \cdot \begin{cases} \pi^i(p^i, (\boldsymbol{p}^C)^{-i}) + \delta_i v^f_{i,(p^i,(\boldsymbol{p}^C)^{-i})} & \text{if } \boldsymbol{p}_0 = \boldsymbol{p}^C, \\ \pi^i(p^i, (\boldsymbol{p}^*)^{-i}) + \delta_i v^f_{i,(p^i,(\boldsymbol{p}^*)^{-i})} & \text{if } \boldsymbol{p}_0 \neq \boldsymbol{p}^C, \end{cases} \\
&= \frac{1}{1-\delta_i} \cdot \begin{cases} \pi^i(\boldsymbol{p}^C) & \text{if } \boldsymbol{p}_0 = \boldsymbol{p}^C, \\ \pi^i(\boldsymbol{p}^*) & \text{if } \boldsymbol{p}_0 \neq \boldsymbol{p}^C \end{cases} \\
&= v^f_{i,\boldsymbol{p}_0}.
\end{aligned}$$

We thus conclude that $\boldsymbol{v}^f$ is a fixed point of the operator $v_{i,\boldsymbol{p}_0} \mapsto \max_{\sigma^i_1 \in \boldsymbol{\Sigma}^i_1} \boldsymbol{V}_1(\boldsymbol{\sigma}^f_1, \sigma^i_1, \boldsymbol{v})_{i,\boldsymbol{p}_0}$.

Step 3 of Algorithm 1: Applying (A.14), (A.15) and (A.16) in a similar way as in step 2 above, we obtain that

$$\sigma^f_0 \in \operatorname{argmax}_{\tau^i_0 \in \boldsymbol{\Sigma}^i_0} \tilde{V}^i_0((\tau^i_0, \sigma^f_1)|(\boldsymbol{\sigma}^f_0, \boldsymbol{\sigma}^f_1)^{-i}),$$

where

$$\tilde{V}_0^i((\tau_0^i,\sigma_1^f)|(\boldsymbol{\sigma}_0^f,\boldsymbol{\sigma}_1^f)^{-i}) = \sum_{p_0^i\in\mathcal{A}} \tau^i(p_0^i)\left\{\pi^i(p_0^i,(\boldsymbol{p}^C)^{-i}) + \delta_i v^f_{i,(p_0^i,(\boldsymbol{p}^C)^{-i})}\right\}.$$

We thus conclude that $\boldsymbol{\sigma}_0^f$ satisfies (2.15). Lastly, the combination of the above equation with (A.14) yields for each $i \in [n]$,

$$\tilde{V}_0^i(\boldsymbol{\sigma}^f) = \frac{1}{1-\delta_i}\pi^i(\boldsymbol{p}^C). \tag{A.17}$$

□

A.4 Proof of Proposition 3

Recall that $\alpha_t = \alpha \in (0,1]$ for each $t \geq 0$ and $(Q^{i*})_{i=1}^n$ is a fixed point of Algorithm 2. Furthermore, for $\boldsymbol{s} = (s_1,\boldsymbol{p}_0) \in \mathcal{S}\times\mathcal{A}^n$, each firm $i \in [n]$ chooses an action according to (2.24) and consequently

$$\max_{p\in\mathcal{A}} Q^{i*}(\hat{\boldsymbol{s}},p) = Q^{i*}(\hat{\boldsymbol{s}},w^i(\hat{\boldsymbol{s}})).$$

Because $(Q^{i*})_{i=1}^n$ is a fixed point of Algorithm 2, then the next update of Q^{i*} satisfies

$$Q^{i*}(\boldsymbol{s},w^i(\boldsymbol{s})) = (1-\alpha)Q_w^{i*}(\boldsymbol{s},w^i(\boldsymbol{s}))+\alpha\left\{\pi^i(\boldsymbol{w}(\boldsymbol{s}),\boldsymbol{s}) + \delta_i\mathbb{E}_{\hat{\boldsymbol{s}}}\left[\max_{p\in\mathcal{A}} Q^{i*}(\hat{\boldsymbol{s}},p)\right]\right\}, \tag{A.18}$$

where $\hat{s} = (s_2,\boldsymbol{w}(\boldsymbol{s}))$ represents the new state after the firms play with $\boldsymbol{w}(\boldsymbol{s})$. Combining the latter equation with (A.18), using that $\alpha \neq 0$ and $\boldsymbol{s} = (s_1,\boldsymbol{p}_0)$, yields

$$Q^{i*}(\boldsymbol{s},w^i(\boldsymbol{s})) = \pi^i(\boldsymbol{w}(\boldsymbol{s}),s_1) + \delta_i\sum_{s_2\in\mathcal{S}}\mathbb{P}(s_2|\boldsymbol{w}(\boldsymbol{s}),s_1)Q^{i*}(\hat{\boldsymbol{s}},w^i(\hat{\boldsymbol{s}})). \tag{A.19}$$

It follows from Proposition 1 that for each $\boldsymbol{s} = (s_1,\boldsymbol{p}_0) \in \mathcal{S}\times\mathcal{A}^n$

$$Q^{i*}(\boldsymbol{s},w^i(\boldsymbol{s})) = \tilde{V}_1^i(\boldsymbol{s},w^i(\boldsymbol{s})|\boldsymbol{w}^{-i}(\boldsymbol{s})).$$

□

A.5 Proof of Proposition 4

Recall that $\alpha_t = \alpha \in (0,1]$ for each $t \geq 0$, $\boldsymbol{Q}^* = (Q^{i*})_{i=1}^n$ is a fixed point of Algorithm 2, and (2.26) holds for each $i \in [n]$ and $\boldsymbol{s} = (s_1, \boldsymbol{p}_0) \in \mathcal{S} \times \mathcal{A}^n$. We use steps 1 and 2 of Algorithm 1 to show that $\boldsymbol{w} = \{w^i(\boldsymbol{s}) | i \in [n], \boldsymbol{s} \in \mathcal{S} \times \mathcal{A}^n\}$ is a Nash equilibrium from time $t = 1$.

Step 1 of Algorithm 1: We plug $\boldsymbol{w}$ into equation (2.16) and solve it as a linear system with unknowns $v_{i,\boldsymbol{s}}$ for each $(i, \boldsymbol{s}) \in [n] \times \mathcal{S} \times \mathcal{A}^n$ and obtain

$$v_{i,\boldsymbol{s}} = \pi^i(\boldsymbol{w}(\boldsymbol{s}), s_1) + \delta_i \sum_{s_2 \in \mathcal{S}} \mathbb{P}(s_2 | \boldsymbol{w}(\boldsymbol{s}), s_1) v_{i,\hat{s}}, \tag{A.20}$$

where $\hat{s} = (s_2, \boldsymbol{w}(\boldsymbol{s}))$. By Proposition 1, $v_{i,\boldsymbol{s}} = \tilde{V}_1^i(\boldsymbol{s}, w^i(\boldsymbol{s}) | \boldsymbol{w}^{-i}(\boldsymbol{s}))$ for each $\boldsymbol{s} \in \mathcal{S} \times \mathcal{A}^n$, $i \in [n]$. Moreover, by Proposition 3,

$$v_{i,\boldsymbol{s}}^l = Q^{i*}(\boldsymbol{s}, w^i(\boldsymbol{s})). \tag{A.21}$$

Step 2 of Algorithm 1: We plug in $\boldsymbol{v} = (v_{i,\boldsymbol{s}})_{i \in [n], \boldsymbol{s} \in \mathcal{S} \times \mathcal{A}^n}$ and $\boldsymbol{w}$ into (2.17) to show that $\boldsymbol{v}$ is a fixed point of the operator $v_{i,\boldsymbol{s}} \mapsto \max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{w}, \sigma_1^i, \boldsymbol{v})_{i,\boldsymbol{s}}$. By (2.26) and (A.21),

$$\begin{aligned}
&\max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{w}, \sigma_1^i, \boldsymbol{v})_{i,\boldsymbol{s}} = \max_{p_1^i \in \mathcal{A}} \boldsymbol{V}_1(\boldsymbol{w}, p_1^i, \boldsymbol{v})_{i,\boldsymbol{s}} = \boldsymbol{V}_1(\boldsymbol{w}, w^i, \boldsymbol{v})_{i,\boldsymbol{s}} \\
&= \pi^i(\boldsymbol{w}(\boldsymbol{s}), s_1) + \delta_i \sum_{s_2 \in \mathcal{S}} \mathbb{P}(s_2 | \boldsymbol{w}(\boldsymbol{s}), s_1) v_{i,\hat{s}} = v_{i,\boldsymbol{s}}.
\end{aligned}$$

The above verification of the first two steps of Algorithm 1 implies that $\boldsymbol{w} = \{l^i(\boldsymbol{s}) | i \in [n], \boldsymbol{s} \in \mathcal{S} \times \mathcal{A}^n\}$ is a Nash equilibrium from time $t = 1$.

□

A.6 Proof of Theorem 4

We break down the proof of Theorem 4 into two main steps: (I) We prove Lemma 1 below which concludes the first claim of Theorem 4 and also characterizes the values of the Q-function given by (2.22) for each $t \geq T$; (II) We use the latter claim to compute the limit in equation (2.28).

(I) We formulate and establish Lemma 1. It uses the definition $\tilde{\alpha}_k := (1 - \alpha_k(1 - \delta_i))$, for each $k \in \mathbb{N}$, and the convention that $\prod_{k=l}^{l-1} \tilde{\alpha}_k = 1$ for each $l \in \mathbb{N}$.

Lemma 1. *If the assumptions of Theorem 4 hold, then for each $i \in [n]$, $t \geq T$ and $p \in \mathcal{A} \setminus \{p^C\}$, $Q_t^i(\boldsymbol{p}_{t-1}, p^C) > Q_t^i(\boldsymbol{p}_{t-1}, p)$ and $p_t^i = p^C$. Moreover,*

$$Q_{T+1}^i(\boldsymbol{s}, p) = \begin{cases} (1 - \alpha_T) Q_T^i(\boldsymbol{p}_{T-1}, p^C) + \alpha_T[\pi^i(\boldsymbol{p}^C) + \delta_i Q_T^i(\boldsymbol{p}^C, p^C)] & \text{if } (\boldsymbol{s}, p) = (\boldsymbol{p}_{T-1}, p^C), \\ Q_T^i(\boldsymbol{s}, p) & \text{otherwise,} \end{cases} \quad \text{(A.22)}$$

and for each $t \geq T + 1$

$$Q_t^i(\boldsymbol{s}, p) = \begin{cases} \prod_{k=T+1}^{t-1} \tilde{\alpha}_k Q_{T+1}^i(\boldsymbol{p}^C, p^C) + \sum_{k=T+1}^{t-1} \prod_{l=k+1}^{t-1} \tilde{\alpha}_l \alpha_k \pi^i(\boldsymbol{p}^C) & \text{if } (\boldsymbol{s}, p) = (\boldsymbol{p}^C, p^C), \\ Q_{T+1}^i(\boldsymbol{s}, p) & \text{otherwise.} \end{cases} \quad \text{(A.23)}$$

Proof of Lemma 1. We fix $i \in [n]$ and note that Assumption (i) in Theorem 4 trivially implies that for each $p \in \mathcal{A} \setminus \{p^C\}$, $Q_T^i(\boldsymbol{p}_{T-1}, p^C) > Q_T^i(\boldsymbol{p}_{T-1}, p)$. The latter equation and Algorithm 4 imply that $\boldsymbol{s}_{T+1} = \boldsymbol{p}_T = \boldsymbol{p}^C$. We thus easily concluded the first part of the lemma To prove the statements in Lemma 1 with $t \geq T + 1$ we use strong induction.

• *Case* $t = T + 1$. By Lemma 1 with $t = T$, $(\boldsymbol{s}_T, p_T^i) = (\boldsymbol{p}_{T-1}, p^C)$. Using (2.22) for each $(\boldsymbol{s}, p) \neq (\boldsymbol{p}_{T-1}, p^C)$, $Q_{T+1}^i(\boldsymbol{s}, p) = Q_T^i(\boldsymbol{s}, p)$ and

$$Q_{T+1}^i(\boldsymbol{p}_{T-1}, p^C) = (1 - \alpha_T) Q_T^i(\boldsymbol{p}_{T-1}, p^C) + \alpha_T[\pi^i(\boldsymbol{p}^C) + \delta_i \max_{p \in \mathcal{A}} Q_T^i(\boldsymbol{p}^C, p)]. \quad \text{(A.24)}$$

Assumption (i) in Theorem 4 implies that $\max_{p \in \mathcal{A}} Q_T^i(\boldsymbol{p}^C, p) = Q_T^i(\boldsymbol{p}^C, p^C)$, which can be replaced into (A.24) to conclude the proof of (A.22) in Lemma 1. Now, we use (A.22) to show that $Q_{T+1}^i(\boldsymbol{p}_T, p^C) > Q_{T+1}^i(\boldsymbol{p}_T, p)$ for each $p \in \mathcal{A} \setminus \{p^C\}$, where $\boldsymbol{p}_T = \boldsymbol{p}^C$. We do so in two cases. First, assume that $\boldsymbol{p}_{T-1} \neq \boldsymbol{p}^C$. Then, by (A.22), $Q_{T+1}^i(\boldsymbol{p}^C, p) = Q_T^i(\boldsymbol{p}^C, p)$ for each $p \in \mathcal{A}$. The equality $\max_{p \in \mathcal{A}} Q_T^i(\boldsymbol{p}^C, p) = Q_T^i(\boldsymbol{p}^C, p^C)$ shows that $Q_{T+1}^i(\boldsymbol{p}^C, p^C) > Q_{T+1}^i(\boldsymbol{p}^C, p)$. Second, assume that $\boldsymbol{p}_{T-1} = \boldsymbol{p}^C$. Using (A.22) and Assumption (ii) in Theorem 4, we obtain for each $p \in \mathcal{A} \setminus \{p^C\}$

$$\begin{aligned} Q^i_{T+1}(\boldsymbol{p}^C, p^C) &= (1-\alpha_T)Q^i_T(\boldsymbol{p}^C, p^C) + \alpha_T[\pi^i(\boldsymbol{p}^C) + \delta_i Q^i_T(\boldsymbol{p}^C, p^C)] \\ &= (1-\alpha_T+\alpha_T\delta_i)Q^i_T(\boldsymbol{p}^C, p^C) + \alpha_T\pi^i(\boldsymbol{p}^C) \\ &\geq (1-\alpha_T+\alpha_T\delta_i)Q^i_T(\boldsymbol{p}^C, p^C) + \alpha_T(1-\delta_i)Q^i_T(\boldsymbol{p}^C, p) \\ &= (1-\alpha_T(1-\delta_i))\underbrace{[Q^i_T(\boldsymbol{p}^C, p^C) - Q^i_T(\boldsymbol{p}^C, p)]}_{>0\text{, by (i) in Theorem 4}} + Q^i_T(\boldsymbol{p}^C, p). \end{aligned} \tag{A.25}$$

Given that $\alpha_T(1-\delta_i) < 1$, from (A.25) for each $p \in \mathcal{A} \setminus \{p^C\}$, we obtain $Q^i_{T+1}(\boldsymbol{p}^C, p^C) > Q^i_T(\boldsymbol{p}^C, p) = Q^i_{T+1}(\boldsymbol{p}^C, p)$. Algorithm 4 implies that $\boldsymbol{p}_{T+1} = \boldsymbol{p}^C$. Finally, note that for $t = T+1$, (A.23) trivially holds true since $\prod_{k=T+1}^{T} \tilde{\alpha_k} = 1$.

• *Inductive case*. Let $t > T+1$ and assume that Lemma 1 holds true for each $T+1 \leq k \leq t$, we prove it for $t+1$. By the inductive hypothesis, $\boldsymbol{s}_{k+1} = \boldsymbol{p}_k = \boldsymbol{p}^C$ for each $T+1 \leq k \leq t$. Therefore, using $Q^i_t(\boldsymbol{p}^C, p^C) > Q^i_t(\boldsymbol{p}^C, p)$ for each $p \in \mathcal{A} \setminus \{p^C\}$ and rule (2.22) with $(\boldsymbol{s}, p) = (\boldsymbol{s}_t, p^i_t) = (\boldsymbol{p}^C, p^C)$,

$$\begin{aligned} Q^i_{t+1}(\boldsymbol{p}^C, p^C) &= (1-\alpha_t)Q^i_t(\boldsymbol{p}^C, p^C) + \alpha_t\left[\pi^i(\boldsymbol{p}^C) + \delta_i \max_{p\in\mathcal{A}} Q^i_t(\boldsymbol{p}^C, p)\right] \\ &= (1-\alpha_t)Q^i_t(\boldsymbol{p}^C, p^C) + \alpha_t\left[\pi^i(\boldsymbol{p}^C) + \delta_i Q^i_t(\boldsymbol{p}^C, p^C)\right] \\ &= (1-\alpha_t(1-\delta_i))Q^i_t(\boldsymbol{p}^C, p^C) + \alpha_t\pi^i(\boldsymbol{p}^C). \end{aligned} \tag{A.26}$$

Moreover, because $\boldsymbol{p}_k = \boldsymbol{p}^C$ for each $T+1 \leq k \leq t$, by (A.22) and rule (2.22) for each $p \in \mathcal{A} \setminus \{p^C\}$,

$$Q^i_{t+1}(\boldsymbol{p}^C, p) = Q^i_t(\boldsymbol{p}^C, p) = \cdots = Q^i_T(\boldsymbol{p}^C, p). \tag{A.27}$$

Combining (A.26), (A.27) and Assumption (ii) in Theorem 4, we obtain for each $p \in \mathcal{A} \setminus \{p^C\}$

$$\begin{aligned} Q^i_{t+1}(\boldsymbol{p}^C, p^C) &> (1-\alpha_t(1-\delta_i))Q^i_T(\boldsymbol{p}^C, p) + \alpha_t\pi^i(\boldsymbol{p}^C) \\ &\geq (1-\alpha_t(1-\delta_i))Q^i_T(\boldsymbol{p}^C, p) + \alpha_t(1-\delta_i)Q^i_T(\boldsymbol{p}^C, p) \\ &= Q^i_T(\boldsymbol{p}^C, p) = Q^i_{t+1}(\boldsymbol{p}^C, p). \end{aligned} \tag{A.28}$$

It follows that $Q^i_{t+1}(\boldsymbol{p}^C, p^C) > Q^i_{t+1}(\boldsymbol{p}^C, p)$ for each $p \in \mathcal{A} \setminus \{p^C\}$. Finally, since by the inductive hypothesis, (A.23) holds for $T+1 \leq k \leq t$, plugging into (A.26) yields

$$\begin{aligned} Q^i_{t+1}(\boldsymbol{p}^C, p^C) &= (1-\alpha_t(1-\delta_i))Q^i_t(\boldsymbol{p}^C, p^C) + \alpha_t \pi^i(\boldsymbol{p}^C) \\ &= \tilde{\alpha}_t \prod_{k=T+1}^{t-1} \tilde{\alpha}_k Q^i_{T+1}(\boldsymbol{p}^C, p^C) + \tilde{\alpha}_t \sum_{k=T+1}^{t-1} \prod_{l=k+1}^{t-1} \tilde{\alpha}_l \alpha_k \pi^i(\boldsymbol{p}^C) + \alpha_t \pi^i(\boldsymbol{p}^C) \\ &= \prod_{k=T+1}^{t} \tilde{\alpha}_k Q^i_{T+1}(\boldsymbol{p}^C, p^C) + \sum_{k=T+1}^{t} \prod_{l=k+1}^{t} \tilde{\alpha}_l \alpha_k \pi^i(\boldsymbol{p}^C). \end{aligned}$$

Thus, concluding the proof of the claim for $t+1$. □

(II) We use Lemma 1 to compute $Q^{i*}_{\epsilon\to 0}(\boldsymbol{s}, p) := \lim_{t\to\infty} Q^i_t(\boldsymbol{s}, p)$.

(a) Case $(\boldsymbol{s}, p) = (\boldsymbol{p}^C, p^C)$. By (A.23), for each $t \geq T+1$, $i \in [n]$

$$Q^i_{t+1}(\boldsymbol{p}^C, p^C) = \prod_{k=T+1}^{t} \tilde{\alpha}_k Q^i_{T+1}(\boldsymbol{p}^C, p^C) + \sum_{k=T+1}^{t} \prod_{l=k+1}^{t} \tilde{\alpha}_l \alpha_k \pi^i(\boldsymbol{p}^C) \tag{A.29}$$

By definition of $\tilde{\alpha}_k = 1-\alpha_k(1-\delta_i)$, $\tilde{\alpha}_k \in (0,1)$ for each $k \geq 1$. Using Assumption 4, we obtain the following

$$\prod_{k=T+1}^{t} \tilde{\alpha}_k = e^{\sum_{k=T+1}^{t} \log(\tilde{\alpha}_k)} \leq e^{\sum_{k=T+1}^{t} \tilde{\alpha}_k - 1} = e^{-(1-\delta_i)\sum_{k=T+1}^{t} \alpha_k} \to 0 \text{ as } t \to \infty. \tag{A.30}$$

Thus, $\lim_{t\to\infty} \prod_{k=T+1}^{t} \tilde{\alpha}_k = 0$. Combining the latter fact with (A.29) yields

$$Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}^C, p^C) = \lim_{t\to\infty} Q^i_t(\boldsymbol{p}^C, p^C) = \lim_{t\to\infty} \sum_{k=T+1}^{t} \prod_{l=k+1}^{t} \tilde{\alpha}_l \alpha_k \pi^i(\boldsymbol{p}^C) = \alpha(\delta_i)\pi^i(\boldsymbol{p}^C) \quad .$$

(b) Case $(\boldsymbol{s}, p) = (\boldsymbol{p}_{T-1}, p^C)$ and $\boldsymbol{p}_{T-1} \neq \boldsymbol{p}^C$. Using rule (2.22), the Q-values at time $t = T$,

$$Q^i_{T+1}(\boldsymbol{p}_{T-1}, p^C) = (1-\alpha_T)Q^i_T(\boldsymbol{p}_{T-1}, p^C) + \alpha_T \left[\pi^i(\boldsymbol{p}^C) + \delta_i Q^i_T(\boldsymbol{p}^C, p^C)\right].$$

(c) Case $(\boldsymbol{s}, p)$ not covered by (a) and (b) above. From Lemma 1, $Q^i_{t+1}(\boldsymbol{s}, p) = Q^i_T(\boldsymbol{s}, p)$ for each $t \geq T$. Thus, $Q^{i*}_{\epsilon\to 0}(\boldsymbol{s}, p) = Q^i_T(\boldsymbol{s}, p)$.

□

A.7 Proof of Proposition 5

We start by proving that for each $\boldsymbol{s} \in \mathcal{A}^n$

$$\boldsymbol{w}^*_{\epsilon\to 0}(\boldsymbol{s}) = \boldsymbol{p}^C.$$

We split the proof of the latter fact in three cases where either $\boldsymbol{s} = \boldsymbol{p}^C$, or $\boldsymbol{s} = \boldsymbol{p}_{T-1} \neq \boldsymbol{p}^C$, or $\boldsymbol{s} \in \mathcal{A}^n \setminus \{\boldsymbol{p}^C, \boldsymbol{p}_{T-1}\}$. We fix $i \in [n]$ for the entire proof.

- Case $\boldsymbol{s} = \boldsymbol{p}^C$. By (2.28),

$$Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}^C, p) = \begin{cases} \alpha(\delta_i)\pi^i(\boldsymbol{p}^C) & \text{if } p = p^C, \\ Q^i_T(\boldsymbol{p}^C, p) & \text{if } p \neq p^C. \end{cases} \tag{A.31}$$

By Assumption (ii) in Proposition 5 with $\boldsymbol{s} = \boldsymbol{p}^C$, $\pi^i(\boldsymbol{p}^C) \geq (1-\delta_i)Q^i_T(\boldsymbol{p}^C, p)$ for each $p \in \mathcal{A} \setminus \{p^C\}$. Combining the latter inequality with Assumption $\alpha(\delta_i)(1-\delta_i) > 1$, by (A.31), $Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}^C, p^C) > Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}^C, p)$ for each $p \in \mathcal{A} \setminus \{p^C\}$. Thus, $\text{argmax}_{p\in\mathcal{A}} Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}^C, p) = \{p^C\}$, which implies that $w^{i*}_{\epsilon\to 0}(\boldsymbol{p}^C) = p^C$.

- Case $\boldsymbol{s} = \boldsymbol{p}_{T-1} \neq \boldsymbol{p}^C$. By (2.28),

$$Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p) = \begin{cases} (1-\alpha_T)Q^i_T(\boldsymbol{p}_{T-1}, p^C) + \alpha_T\left[\pi^i(\boldsymbol{p}^C) + \delta_i Q^i_T(\boldsymbol{p}^C, p^C)\right] & \text{if } p = p^C, \\ Q^i_T(\boldsymbol{p}_{T-1}, p) & \text{if } p \neq p^C. \end{cases} \tag{A.32}$$

By Assumption (ii) in Proposition 5 with $\boldsymbol{s} = \boldsymbol{p}_{T-1}$, $\pi^i(\boldsymbol{p}^C) \geq Q^i_T(\boldsymbol{p}_{T-1}, p) - \delta_i Q^i_T(\boldsymbol{p}^C, p)$ for each $p \in \mathcal{A} \setminus \{p^C\}$. Thus, for each $p \in \mathcal{A} \setminus \{p^C\}$

$$\begin{aligned} &Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p^C) \\ &\geq (1-\alpha_T)Q^i_T(\boldsymbol{p}_{T-1}, p^C) + \alpha_T\left[Q^i_T(\boldsymbol{p}_{T-1}, p) - \delta_i Q^i_T(\boldsymbol{p}^C, p) + \delta_i Q^i_T(\boldsymbol{p}^C, p^C)\right] \\ &= (1-\alpha_T)\underbrace{\left[Q^i_T(\boldsymbol{p}_{T-1}, p^C) - Q^i_T(\boldsymbol{p}_{T-1}, p)\right]}_{>0,\text{ by (i) in Proposition 5}} + \alpha_T\delta_i \underbrace{\left[Q^i_T(\boldsymbol{p}^C, p^C) - Q^i_T(\boldsymbol{p}^C, p)\right]}_{>0,\text{ by (i) in Proposition 5}} + Q^i_T(\boldsymbol{p}_{T-1}, p) \end{aligned} \tag{A.33}$$

From (A.33), $Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p^C) > Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p)$ for each $p \in \mathcal{A} \setminus \{p^C\}$. It follows that $w^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}) = p^C$.

• Case $\boldsymbol{s} \in \mathcal{A}^n \setminus \{\boldsymbol{p}^C, \boldsymbol{p}_{T-1}\}$. By (2.28), $Q^{i*}_{\epsilon\to 0}(\boldsymbol{s}, p) = Q^i_T(\boldsymbol{s}, p)$ for each $p \in \mathcal{A}$. By Assumption (i) in Proposition 5, $Q^i_T(\boldsymbol{s}, p^C) > Q^i_T(\boldsymbol{s}, p)$ for each $p \in \mathcal{A} \setminus \{p^C\}$. It follows that $w^{i*}_{\epsilon\to 0}(\boldsymbol{s}) = p^C$.

Finally, we prove that $\boldsymbol{w}^*_{\epsilon\to 0}$ is a Nash equilibrium from time $t = 1$ if and only if $\boldsymbol{p}^C$ is a Nash equilibrium of the one-stage game $(\pi^i(\cdot))^n_{i=1}$.

⋄ Suppose that $\boldsymbol{p}^C$ is a Nash equilibrium of the one-stage game $(\pi^i(\cdot))^n_{i=1}$. We use Algorithm 1 to show that $\boldsymbol{w}^*_{\epsilon\to 0}$ is a Nash equilibrium from time $t = 1$. By (i) in Algorithm 1, we first plug $\boldsymbol{w}^*_{\epsilon\to 0} = \boldsymbol{p}^C$ into equation (2.16) and solve it as a linear system with unknowns $\boldsymbol{v} = (v_{i,\boldsymbol{p}_0})_{\boldsymbol{p}_0 \in \mathcal{A}^n}$, as follows:

$$\begin{aligned} v_{i,\boldsymbol{p}_0} &= \boldsymbol{V}_1(\boldsymbol{w}^*_{\epsilon\to 0}, \boldsymbol{w}^*_{\epsilon\to 0}, \boldsymbol{v})_{i,\boldsymbol{p}_0} \\ &\underbrace{=}_{\text{By (2.10)}} \pi^i(\boldsymbol{p}^C) + \delta_i v_{i,\boldsymbol{p}^C} \end{aligned} \tag{A.34}$$

Solving (A.34) for $\boldsymbol{v}$, yields for each $\boldsymbol{p}_0 \in \mathcal{A}^n$

$$v_{i,\boldsymbol{p}_0} = \frac{1}{1-\delta_i}\pi^i(\boldsymbol{p}^C). \tag{A.35}$$

Following Algorithm 1 part (ii), we plug in $\boldsymbol{v}$ and $\boldsymbol{w}^*_{\epsilon\to 0} = \boldsymbol{p}^C$ into (2.17) to check if $\boldsymbol{v}$ is a fixed point of the operator $v_{i,\boldsymbol{p}_0} \mapsto \max_{\sigma^i_1 \in \boldsymbol{\Sigma}^i_1} \boldsymbol{V}_1(\boldsymbol{w}^*_{\epsilon\to 0}, \sigma^i_1, \boldsymbol{v})_{i,\boldsymbol{p}_0}$. Indeed, by (2.10) and (A.35),

$$\begin{aligned} \max_{\sigma^i_1 \in \boldsymbol{\Sigma}^i_1} \boldsymbol{V}_1(\boldsymbol{w}^*_{\epsilon\to 0}, \sigma^i_1, \boldsymbol{v})_{i,\boldsymbol{p}_0} &= \max_{\sigma^i_1 \in \boldsymbol{\Sigma}^i_1} \sum_{p^i_1 \in \mathcal{A}} \sigma^i_1(p^i_1|\boldsymbol{p}_0)[\pi^i(p^i_1, (\boldsymbol{p}^C)^{-i}) + \delta_i v_{i,(p^i_1,(\boldsymbol{p}^C)^{-i})}] \\ &= \max_{\sigma^i_1 \in \boldsymbol{\Sigma}^i_1} \sum_{p^i_1 \in \mathcal{A}} \sigma^i_1(p^i_1|\boldsymbol{p}_0)\left[\pi^i(p^i_1, (\boldsymbol{p}^C)^{-i}) + \frac{\delta_i}{1-\delta_i}\pi^i(\boldsymbol{p}^C)\right]. \end{aligned} \tag{A.36}$$

Since $\boldsymbol{p}^C$ is a Nash equilibrium of the one-stage game $(\pi^i(\cdot))^n_{i=1}$, the maximum in (A.36) is achieved at $\sigma_1^i(p_1^i|\boldsymbol{p}_0) = p^C$ for each $p_1^i \in \mathcal{A}$. Thus,

$$\max_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{w}^*_{\epsilon\to 0}, \sigma_1^i, \boldsymbol{v})_{i,\boldsymbol{p}_0} = \frac{1}{1-\delta_i}\pi^i(\boldsymbol{p}^C) = v_{i,\boldsymbol{p}_0}.$$

By Algorithm 1, $\boldsymbol{w}^*_{\epsilon\to 0}$ is a Nash equilibrium from time $t = 1$.

$\diamond$ Suppose that $\boldsymbol{w}^*_{\epsilon\to 0} = \boldsymbol{p}^C$ is a Nash equilibrium from time $t = 1$. By definition (2.14), for each $\boldsymbol{p}_0 \in \mathcal{A}^n$

$$\boldsymbol{w}^{i*}_{\epsilon\to 0}(\boldsymbol{p}_0) = p^C \in \text{argmax}_{\sigma_1^i \in \boldsymbol{\Sigma}_1^i} \tilde{V}_1^i(\boldsymbol{p}_0, \sigma_1^i | \boldsymbol{w}^{-i*}_{\epsilon\to 0}).$$

By the above and equation (2.9), for each $\boldsymbol{p}_0 \in \mathcal{A}^n$ and $\sigma_1^i \in \boldsymbol{\Sigma}_1^i$

$$\sum_{t=1}^{\infty} \delta_i^{t-1}\pi^i(\boldsymbol{p}^C) \geq \mathbb{E}_{(\sigma_1^i, \boldsymbol{w}^{-i*}_{\epsilon\to 0})}\left[\sum_{t=1}^{\infty} \delta_i^{t-1}\pi^i(\boldsymbol{p}_t)\middle|\boldsymbol{p}_0\right]. \tag{A.37}$$

For each $\hat{p} \in \mathcal{A} \setminus \{p^C\}$, define $\hat{\sigma}_1^i$ as follows: $\hat{\sigma}_1^i(p|\boldsymbol{p}^*) = 1$ if $p = \hat{p}$, and $\hat{\sigma}_1^i(p|\boldsymbol{p}^*) = 0$ if $p \neq \hat{p}$. Moreover, let $\hat{\sigma}_1^i(\cdot|\boldsymbol{p}_0) = p^C$ for any $\boldsymbol{p}_0 \neq \boldsymbol{p}^*$. Taking $\boldsymbol{p}_0 = \boldsymbol{p}^*$ and $\sigma_1^i = \hat{\sigma}_1^i$ in (A.37) yields,

$$\begin{aligned}
\frac{1}{1-\delta_i}\pi^i(\boldsymbol{p}^C) &\geq \mathbb{E}_{\hat{\sigma}_1^i}\left[\pi^i(p_1^i, (\boldsymbol{p}^C)^{-i}) + \sum_{t=2}^{\infty} \delta_i^{t-1}\pi^i(p_t^i, (\boldsymbol{p}^C)^{-i})\middle|\boldsymbol{p}^*\right] \\
&= \pi^i(\hat{p}, (\boldsymbol{p}^C)^{-i}) + \mathbb{E}_{\hat{\sigma}_1^i}\left[\sum_{t=2}^{\infty} \delta_i^{t-1}\pi^i(p_t^i, (\boldsymbol{p}^C)^{-i})\middle|(\hat{p}, (\boldsymbol{p}^C)^{-i})\right] \\
&= \pi^i(\hat{p}, (\boldsymbol{p}^C)^{-i}) + \frac{\delta_i}{1-\delta_i}\pi^i(\boldsymbol{p}^C).
\end{aligned}$$

The above inequality holds for each $\hat{p} \in \mathcal{A} \setminus \{p^C\}$ and $i \in [n]$, implying that $\boldsymbol{p}^C$ is a Nash equilibrium of the one-stage game $(\pi^i(\cdot))_{i=1}^n$.

□

A.8 Proof of Proposition 6

We start by proving that

$$\boldsymbol{w}^*_{\epsilon\to 0}(\boldsymbol{s}) = \begin{cases} \boldsymbol{p}^C & \boldsymbol{s} = \boldsymbol{p}^C, \\ \boldsymbol{p}^* & \boldsymbol{s} \neq \boldsymbol{p}^C. \end{cases}$$

We split the proof of the latter fact in three cases where either $\boldsymbol{s} = \boldsymbol{p}^C$, or $\boldsymbol{s} = \boldsymbol{p}_{T-1} \neq \boldsymbol{p}^C$, or $\boldsymbol{s} \in \mathcal{A}^n \setminus \{\boldsymbol{p}^C, \boldsymbol{p}_{T-1}\}$. We fix $i \in [n]$ for the entire proof.

• Case $\boldsymbol{s} = \boldsymbol{p}^C$. This case is identical to case $\boldsymbol{s} = \boldsymbol{p}^C$ in the Proof of Proposition 5, so we omit it. However, we recall that this case uses the assumptions $\alpha(\delta_i)(1-\delta_i) > 1$ and Assumption (ii) in Proposition 6. Thus, $w^{i*}_{\epsilon\to 0}(\boldsymbol{p}^C) = p^C$.

• Case $\boldsymbol{s} = \boldsymbol{p}_{T-1} \neq \boldsymbol{p}^C$. By (2.28),

$$Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p) = \begin{cases} (1-\alpha_T) Q^i_T(\boldsymbol{p}_{T-1}, p^C) + \alpha_T \left[\pi^i(\boldsymbol{p}^C) + \delta_i Q^i_T(\boldsymbol{p}^C, p^C)\right] & \text{if } p = p^C, \\ Q^i_T(\boldsymbol{p}_{T-1}, p) & \text{if } p \neq p^C. \end{cases} \tag{A.38}$$

By Assumption (i) in Proposition 6, $Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p^*) > Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p)$ for each $p \in \mathcal{A} \setminus \{p^*\}$. Thus, $w^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}) = p^*$.

• Case $\boldsymbol{s} \in \mathcal{A}^n \setminus \{\boldsymbol{p}^C, \boldsymbol{p}_{T-1}\}$. By (2.28), $Q^{i*}_{\epsilon\to 0}(\boldsymbol{s}, p) = Q^i_T(\boldsymbol{s}, p)$ for each $p \in \mathcal{A}$. By Assumption (i) in Proposition 6, $Q^i_T(\boldsymbol{s}, p^*) > Q^i_T(\boldsymbol{s}, p)$ for each $p \in \mathcal{A} \setminus \{p^*\}$. It follows that $w^{i*}_{\epsilon\to 0}(\boldsymbol{s}) = p^*$.

Finally, by Proposition 2, we know that under Assumption 3-(ii), $\boldsymbol{w}^*_{\epsilon\to 0}$ is a Nash equilibrium from time $t = 1$, since $\boldsymbol{w}^*_{\epsilon\to 0} = \boldsymbol{\sigma}^f_1$.

□

A.9 Proof of Proposition 7

We start by proving that

$$\boldsymbol{w}^*_{\epsilon\to 0}(\boldsymbol{s}) = \begin{cases} \boldsymbol{p}^C & \boldsymbol{s} = \boldsymbol{p}^C, \\ \boldsymbol{p}^{l+1} & \boldsymbol{s} = \boldsymbol{p}^l, \\ \boldsymbol{p}^* & \boldsymbol{s} \notin \{p^l\}_{l=0}^{k+1}. \end{cases}$$

We split the proof of the latter fact in three cases where either $\boldsymbol{s} = \boldsymbol{p}^C$, or $\boldsymbol{s} = \boldsymbol{p}^j$ for some $j \in [k]$, or $\boldsymbol{s} \in \mathcal{A}^n \setminus \{\boldsymbol{p}^l\}_{l=0}^{k+1}$. We fix $i \in [n]$ for the entire proof.

• Case $\boldsymbol{s} = \boldsymbol{p}^C$. This case is identical to case $\boldsymbol{s} = \boldsymbol{p}^C$ in the Proof of Proposition 5, so we omit it. However, we recall that this case uses the assumptions $\alpha(\delta_i)(1-\delta_i) > 1$ and Assumption (i) in Proposition 7. Thus, $w^{i*}_{\epsilon\to 0}(\boldsymbol{p}^C) = p^C$.

• Case $\boldsymbol{s} = \boldsymbol{p}^j$ for some $j \in [k]$. By Assumption 5, $\boldsymbol{p}_{T-1} \notin \{p^l\}_{l=0}^{k+1}$. By (2.28), $Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}^j, p) = Q^i_T(\boldsymbol{p}^j, p)$ for each $p \in \mathcal{A}$. By Assumption 5-(i), $Q^i_T(\boldsymbol{p}^j, p^{j+1}) > Q^i_T(\boldsymbol{p}^j, p)$ for each $p \in \mathcal{A} \setminus \{p^{j+1}\}$. It follows that $w^{i*}_{\epsilon\to 0}(\boldsymbol{p}^j) = p^{j+1}$.

• Case $\boldsymbol{s} \in \mathcal{A}^n \setminus \{\boldsymbol{p}^l\}_{l=0}^{k+1}$. Since $\boldsymbol{p}_{T-1} \notin \{p^l\}_{l=0}^{k+1}$, by (2.28),

$$Q^{i*}_{\epsilon\to 0} = \begin{cases} Q^{i*}_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p^C) & (\boldsymbol{s}, p) = (\boldsymbol{p}_{T-1}, p^C), \\ Q^i_T(\boldsymbol{s}, p) & (\boldsymbol{s}, p) \neq (\boldsymbol{p}_{T-1}, p^C). \end{cases}$$

By Assumption 5-(ii), $Q^i_T(\boldsymbol{s}, p^*) > \max\{Q^i_T(\boldsymbol{s}, p), Q^i_{\epsilon\to 0}(\boldsymbol{p}_{T-1}, p^C)\}$ for each $p \in \mathcal{A}\setminus\{p^*\}$ and $\boldsymbol{s} \in \mathcal{A} \setminus \{p^l\}_{l=0}^{k+1}$ with $(\boldsymbol{s}, p) \neq (\boldsymbol{p}_{T-1}, p^C)$. It follows that $w^{i*}_{\epsilon\to 0}(\boldsymbol{s}) = p^*$.

□

Example of a Sequence satisfying Assumption 4

For each $k \geq 1$, we let $a_k := \prod_{l=k+1}^{\infty}(1 - \alpha_l(1-\delta_i))\alpha_k$. Suppose that α_k is chosen so that $a_k = \delta_i^{k-1}$. Then,

$$\delta_i = \frac{a_k}{a_{k-1}} = \frac{\alpha_k}{(1 - \alpha_k(1-\delta_i))\alpha_{k-1}}.$$

It follows that $\delta_i(1 - \alpha_k(1-\delta_i))\alpha_{k-1} = \alpha_k$ if and only if

$$\alpha_k = \frac{\delta_i \alpha_{k-1}}{1 + \delta_i(1-\delta_i)\alpha_{k-1}}.$$

With this choice of α_k,

$$\lim_{t\to\infty} Q_t^i(\boldsymbol{p}^C, p^C) = \sum_{k=1}^{\infty} \delta_i^{k-1} \pi^i(\boldsymbol{p}^C) = \frac{1}{1-\delta_i} \pi^i(\boldsymbol{p}^C).$$

Note that if $\alpha_1 \in [0,1)$. Then, $\alpha_2 = \frac{\delta_i \alpha_1}{1+\delta_i(1-\delta_i)\alpha_1} < 1$ if and only if $\delta_i^2 \alpha_1 < 1$. By induction, $\alpha_k < 1$. On the other hand, by definition,

$$\alpha_k > \frac{(1-\delta_i)\delta_i \alpha_{k-1}}{1 + \delta_i(1-\delta_i)\alpha_{k-1}} > \frac{(1-\delta_i)\delta_i \alpha_{k-1}}{2\delta_i(1-\delta_i)\alpha_{k-1}} = \frac{1}{2}.$$

□

A.10 Rewriting the Proof of Fink's Theorem

We rewrite the proof of Theorem 2 of Fink (1964), that is, Theorem 1 in this work. The rewritten proof uses our notation and adds many missing details. We find it necessary to refer to the rewritten proof when establishing the theories of Sections 2.3.1 and 2.4. Section Appendix A.10.1 first proves Proposition 1 and Section Appendix A.10.2 establishes several other propositions and then concludes the proof of Theorem 1.

A.10.1 Proof of Proposition 1

Let $\boldsymbol{\sigma}_1 = (\sigma_1^i, \boldsymbol{\sigma}_1^{-i}) \in \boldsymbol{\Sigma}_1$, $s_1 \in \mathcal{S}$ and $\boldsymbol{p}_0 \in \mathcal{A}^n$ be given. From (2.9),

$$\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i|\boldsymbol{\sigma}_1^{-i}) = \sum_{\boldsymbol{p}_1 \in \mathcal{A}^n} \sigma_1(\boldsymbol{p}_1|\boldsymbol{p}_0, s_1) \left\{ \pi^i(\boldsymbol{p}_1, s_1) + \delta_i \sum_{s_2 \in \mathcal{S}} \mathbb{P}(s_2|\boldsymbol{p}_1, s_1) \mathbb{E}_{\boldsymbol{\sigma}_1, \mathbb{P}} \left[\sum_{t=2}^{\infty} \delta_i^{t-2} r^i(t) \middle| \boldsymbol{p}_1, s_2 \right] \right\}. \tag{A.39}$$

To obtain (2.11) from (A.39), note that the payoff function π^i is time independent, which implies that

$$\mathbb{E}_{\boldsymbol{\sigma}_1, \mathbb{P}} \left[\sum_{t=2}^{\infty} \delta_i^{t-2} r^i(t) \middle| \boldsymbol{p}_1, s_2 \right] = \tilde{V}_1^i(s_2, \boldsymbol{p}_1, \sigma_1^i|\boldsymbol{\sigma}_1^{-i}).$$

We now show that there exists a unique solution to (2.11). Expanding (2.11), for each $s_1 \in \mathcal{S}$ and $\boldsymbol{p}_0 \in \mathcal{A}^n$, we obtain the following

$$\begin{aligned} &\tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i|\boldsymbol{\sigma}_1^{-i}) \\ &= \mathbb{E}_{\boldsymbol{\sigma}_1} \left[\pi^i | \boldsymbol{p}_0, s_1 \right] + \delta_i \sum_{\boldsymbol{p}_1 \in \mathcal{A}^n} \sigma_1(\boldsymbol{p}_1|\boldsymbol{p}_0, s_1) \sum_{s_2 \in \mathcal{S}} \mathbb{P}(s_2|\boldsymbol{p}_1, s_1) \tilde{V}_1^i(s_2, \boldsymbol{p}_1, \sigma_1^i|\boldsymbol{\sigma}_1^{-i}), \end{aligned}$$

which can be rewritten as

$$\begin{aligned} &[1 - \delta_i \sigma_1(\boldsymbol{p}_0|\boldsymbol{p}_0, s_1) \mathbb{P}(s_1|\boldsymbol{p}_0, s_1)] \, \tilde{V}_1^i(s_1, \boldsymbol{p}_0, \sigma_1^i|\boldsymbol{\sigma}_1^{-i}) \\ &- \delta_i \sigma_1(\boldsymbol{p}_0|\boldsymbol{p}_0, s_1) \sum_{s_2 \neq s_1} \mathbb{P}(s_2|\boldsymbol{p}_0, s_1) \tilde{V}_1^i(s_2, \boldsymbol{p}_0, \sigma_1^i|\boldsymbol{\sigma}_1^{-i}) \\ &- \delta_i \sum_{\boldsymbol{p}_1 \neq \boldsymbol{p}_0} \sigma_1(\boldsymbol{p}_1|\boldsymbol{p}_0, s_1) \sum_{s_2 \in \mathcal{S}} \mathbb{P}(s_2|\boldsymbol{p}_1, s_1) \tilde{V}_1^i(s_2, \boldsymbol{p}_1, \sigma_1^i|\boldsymbol{\sigma}_1^{-i}) = \mathbb{E}_{\boldsymbol{\sigma}_1} \left[\pi^i | \boldsymbol{p}_0, s_1 \right]. \end{aligned} \tag{A.40}$$

Let $\mathbb{E}_{\boldsymbol{\sigma}_1}[\pi^i] := (\mathbb{E}_{\boldsymbol{\sigma}_1}[\pi^i|\boldsymbol{p}^1, s^1], \cdots, \mathbb{E}_{\boldsymbol{\sigma}_1}[\pi^i|\boldsymbol{p}^M, s^r])^T \in \mathbb{R}^{rM}$. By (A.40), the vector $\tilde{V}_1^i(\sigma_1^i|\boldsymbol{\sigma}_1^{-i})$ given by

$$\tilde{V}_1^i(\sigma_1^i|\boldsymbol{\sigma}_1^{-i}) := (\tilde{V}_1^i(s^1, \boldsymbol{p}^1, \sigma_1^i|\boldsymbol{\sigma}_1^{-i}), \cdots, \tilde{V}_1^i(s^r, \boldsymbol{p}^M, \sigma_1^i|\boldsymbol{\sigma}_1^{-i}))^T \in \mathbb{R}^{rM} \tag{A.41}$$

satisfies the following linear system

$$\boldsymbol{A}\tilde{V}_1^i(\sigma_1^i|\boldsymbol{\sigma}_1^{-i}) = \mathbb{E}_{\boldsymbol{\sigma}_1}[\pi^i], \tag{A.42}$$

where $\boldsymbol{A}$ is a matrix whose rows and columns are indexed by the set $\mathcal{S} \times \mathcal{A}^n$: the entry in row $(s^j, \boldsymbol{p}^k)$ and column $(s^l, \boldsymbol{p}^o)$ is given by

$$\boldsymbol{A}\left((s^j, \boldsymbol{p}^k), (s^l, \boldsymbol{p}^o)\right) = \begin{cases} 1 - \delta_i \sigma_1(\boldsymbol{p}^k|\boldsymbol{p}^k, s^j)\mathbb{P}(s^j|\boldsymbol{p}^k, s^j) & \text{if } (s^j, \boldsymbol{p}^k) = (s^l, \boldsymbol{p}^o) \\ -\delta_i \sigma_1(\boldsymbol{p}^o|\boldsymbol{p}^k, s^j)\mathbb{P}(s^l|\boldsymbol{p}^o, s^j) & \text{if } (s^j, \boldsymbol{p}^k) \neq (s^l, \boldsymbol{p}^o) \end{cases}. \tag{A.43}$$

For each $(s^j, \boldsymbol{p}^k) \in \mathcal{S} \times \mathcal{A}^n$, the following holds true

$$\begin{aligned} &\boldsymbol{A}\left((s^j, \boldsymbol{p}^k), (s^j, \boldsymbol{p}^k)\right) - \sum_{(s^l, \boldsymbol{p}^o) \neq (s^j, \boldsymbol{p}^k)} \left|\boldsymbol{A}\left((s^j, \boldsymbol{p}^k), (s^l, \boldsymbol{p}^o)\right)\right| \\ &= 1 - \delta_i \sigma_1(\boldsymbol{p}^k|\boldsymbol{p}^k, s^j)\mathbb{P}(s^j|\boldsymbol{p}^k, s^j) - \sum_{(s^l, \boldsymbol{p}^o) \neq (s^j, \boldsymbol{p}^k)} \delta_i \sigma_1(\boldsymbol{p}^o|\boldsymbol{p}^k, s^j)\mathbb{P}(s^l|\boldsymbol{p}^o, s^j) \\ &= 1 - \delta_i \sum_{(s^l, \boldsymbol{p}^o)} \sigma_1(\boldsymbol{p}^o|\boldsymbol{p}^k, s^j)\mathbb{P}(s^l|\boldsymbol{p}^o, s^j) \\ &= 1 - \delta_i \sum_{\boldsymbol{p}^o} \sigma_1(\boldsymbol{p}^o|\boldsymbol{p}^k, s^j) \sum_{s^l} \mathbb{P}(s^l|\boldsymbol{p}^o, s^j) = 1 - \delta_i \end{aligned}. \tag{A.44}$$

From Gershgorin Circle Theorem (See page 244 in Bhatia (2013)), for any eigenvalue of $\boldsymbol{A}$, say λ, there exists $(s^j, \boldsymbol{p}^k) \in \mathcal{S} \times \mathcal{A}^n$ such that

$$|\lambda - \boldsymbol{A}\left((s^j, \boldsymbol{p}^k), (s^j, \boldsymbol{p}^k)\right)| \leq \sum_{(s^l, \boldsymbol{p}^o) \neq (s^j, \boldsymbol{p}^k)} \left|\boldsymbol{A}\left((s^j, \boldsymbol{p}^k), (s^l, \boldsymbol{p}^o)\right)\right|.$$

The above inequality combined with the reverse triangle inequality and equation (A.44), imply that $|\lambda| \geq 1 - \delta_i > 0$. Thus, 0 is not an eigenvalue of $\boldsymbol{A}$ and $\boldsymbol{A}^{-1}$ exists. Therefore, (A.42) has a unique solution. □

A.10.2 Proof of Theorem 1

Before getting into the details of the proof. We summarize some the crucial steps in the proof of Fink (1964):

1. $\boldsymbol{V}_1$ is continuous in its domain of definition (see Proposition 23).

2. For each $\boldsymbol{v} \in \mathbb{R}^{nrM}$ and $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1$, there is a well-defined mapping $(\boldsymbol{v}, \boldsymbol{\sigma}_1) \mapsto T(\boldsymbol{v}, \boldsymbol{\sigma}_1)$ whose $(i, s_1, \boldsymbol{p}_0)$-coordinate is given by

$$T(\boldsymbol{v}, \boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} = \max_{\tau_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}.$$

   The mapping $\boldsymbol{v} \mapsto T(\boldsymbol{v}, \boldsymbol{\sigma}_1)$ is a contraction from $\mathbb{R}^{nrM}$ to itself (see Proposition 24). Thus, there is a well-defined mapping $\boldsymbol{\sigma}_1 \mapsto b(\boldsymbol{\sigma}_1) \in \mathbb{R}^{nrM}$, where $b(\boldsymbol{\sigma}_1)$ is the unique fixed point of $T(\cdot, \boldsymbol{\sigma}_1)$.

3. The set-valued mapping $\Gamma : \boldsymbol{\Sigma}_1 \to 2^{\boldsymbol{\Sigma}_1}$ given by $\boldsymbol{\sigma}_1 \mapsto \Gamma(\boldsymbol{\sigma}_1)$, where

$$\Gamma(\boldsymbol{\sigma}_1) := \{\boldsymbol{\tau}_1 \in \boldsymbol{\Sigma}_1 | b(\boldsymbol{\sigma}_1) = \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, b(\boldsymbol{\sigma}_1))\},$$

   satisfies the hypotheses of Kakutani's theorem (see Theorem 5 and Proposition 25). Therefore, Γ has a fixed point $\boldsymbol{\sigma}_1^* \in \boldsymbol{\Sigma}_1$, i.e., there is a policy in $\boldsymbol{\Sigma}_1$ such that $\boldsymbol{\sigma}_1^* \in \Gamma(\boldsymbol{\sigma}_1^*)$. Such policy is the stationary point of Theorem 1. Moreover, the vector $\boldsymbol{v}^*$ from Theorem 1 is given by $\boldsymbol{v}^* = b(\boldsymbol{\sigma}_1^*)$.

Preliminary Results and Definitions for the Proof of Theorem 1.

Given two nonempty sets X and Y, a correspondence from X to Y is a map $\Gamma : X \longrightarrow 2^Y$ such that for each $x \in X$, $\Gamma(x) \neq \emptyset$. We say that Γ is a self-correspondence on X, if Γ is a correspondence from X to X. If $Y \subset \mathbb{R}^d$ and $\Gamma(x)$ is convex for each $x \in X$, then we say that Γ is convex-valued. Let X and Y be two metric spaces, Γ is said to be closed-valued if $\Gamma(x)$ is a closed subset of Y. Now, Γ is said to be closed at $x \in X$, if for any two sequences $(x_k)_k \subset X$ and $(y_k)_k \subset Y$ with $x_k \to x$ and $y_k \to y \in Y$, if $y_k \in \Gamma(x_k)$ for each k, then $y \in \Gamma(x)$. Moreover, Γ has a closed graph if it is closed at every $x \in X$.

Theorem 5 (Kakutani's Fixed Point Theorem)**.** *Let $X \subset \mathbb{R}^d$ be a nonempty, compact and convex set. If Γ is a convex-valued self-correspondence on X that has a closed graph, then Γ has a fixed point, i.e., there exists $x \in X$ with $x \in \Gamma(x)$.*

For a proof of Kakutani's fixed point theorem see Page 331 in Ok (2007). Proposition 23, Proposition 24 and Proposition 25 below ensure that we can use Kakutani's fixed point theorem to prove Theorem 2.

Proposition 23 (Properties of $\boldsymbol{V}_1$)**.** *The function $\boldsymbol{V}_1$ as given by (2.10) satisfies all of the following:*

(a) $\boldsymbol{V}_1$ is continuous on $\boldsymbol{\Sigma}_1 \times \boldsymbol{\Sigma}_1 \times \mathbb{R}^{nrM}$;

(b) Let $\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1 \in \boldsymbol{\Sigma}_1$ and $\delta := \max_{i\in[n]} \delta_i$. For each $\boldsymbol{v}, \boldsymbol{u} \in \mathbb{R}^{nrM}$, and each $(i, s_1, \boldsymbol{p}_0)$-coordinate

$$\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0} - \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{u})_{i,s_1,\boldsymbol{p}_0} \leq \delta|\boldsymbol{v} - \boldsymbol{u}|_\infty,$$

where $|\cdot|_\infty$ denotes the infinity norm in $\mathbb{R}^{nrM}$;

(c) $\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, \boldsymbol{v})$ is linear in $\boldsymbol{\tau}_1$.

Proof of Proposition 23. Let $\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1 \in \boldsymbol{\Sigma}_1$ and $\boldsymbol{v} \in \mathbb{R}^{nrM}$. From (2.10), for each $(i, s_1, \boldsymbol{p}_0)$-coordinate

$$\begin{aligned} &\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0} \\ &= \sum_{\boldsymbol{p}_1\in\mathcal{A}^n} \tau_1^i(p_1^i|\boldsymbol{p}_0, s_1)\sigma_1^{-i}(\boldsymbol{p}_1^{-i}|\boldsymbol{p}_0, s_1) \left[\pi^i(\boldsymbol{p}_1, s_1) + \delta_i \sum_{s_2\in\mathcal{S}} \mathbb{P}(s_2|\boldsymbol{p}_1, s_1) v_{i,s_2,\boldsymbol{p}_1}\right]. \end{aligned} \tag{A.45}$$

From (A.45), it is straightforward to see that $\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}$ is continuous w.r.t $(\tau_1^i, \boldsymbol{\sigma}_1^{-i})$, and continuous w.r.t. $v_{i,s_2,\boldsymbol{p}_1}$ for all $(i, s_2, \boldsymbol{p}_1)$. Similarly, from (A.45) it is not difficult to see that $\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}$ is linear w.r.t. τ_1^i. Thus, proving (a) and (c). For (b), we estimate

$$\begin{aligned} &\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0} - \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{u})_{i,s_1,\boldsymbol{p}_0} \\ &= \delta_i \sum_{\boldsymbol{p}_1\in\mathcal{A}^n} \tau_1^i(p_1^i|\boldsymbol{p}_0, s_1)\sigma_1^{-i}(\boldsymbol{p}_1|\boldsymbol{p}_0, s_1) \sum_{s_2\in\mathcal{S}} \mathbb{P}(s_2|\boldsymbol{p}_1, s_1)[v_{i,s_2,\boldsymbol{p}_1} - u_{i,s_2,\boldsymbol{p}_1}] \\ &\leq \delta \max_{j,s_2,\boldsymbol{p}_1} |v_{j,s_2,\boldsymbol{p}_1} - u_{j,s_2,\boldsymbol{p}_1}|. \end{aligned}$$

□

The T mapping: From (2.4), we know that $\boldsymbol{\Sigma}_1^i$ is a compact subset of $\mathbb{R}^{(m+1)rM}$. By Proposition 23, $\boldsymbol{V}_1$ is a continuous function. Based on these two observations, it makes sense to define the following mapping:

$$T : \mathbb{R}^{nrM} \times \boldsymbol{\Sigma}_1 \longrightarrow \mathbb{R}^{nrM} \text{ s.t. } (\boldsymbol{v}, \boldsymbol{\sigma}_1) \mapsto T(\boldsymbol{v}, \boldsymbol{\sigma}_1)$$

where the $(i, s_1, \boldsymbol{p}_0)$-coordinate of $T(\boldsymbol{v}, \boldsymbol{\sigma}_1)$ is given by

$$T(\boldsymbol{v}, \boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} := \max_{\tau_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}. \tag{A.46}$$

Proposition 24 (Properties of T). *(i) For each $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1$, the mapping from $\mathbb{R}^{nrM}$ to $\mathbb{R}^{nrM}$ given by $\boldsymbol{v} \mapsto T(\boldsymbol{v}, \boldsymbol{\sigma}_1)$ is a contraction mapping. In particular, for every $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1$, $T(\cdot, \boldsymbol{\sigma}_1)$ has a unique fixed point.*

(ii) For each $\boldsymbol{v} \in \mathbb{R}^{nrM}$, the mapping from $\boldsymbol{\Sigma}_1$ to $\mathbb{R}^{nrM}$ given by $\boldsymbol{\sigma}_1 \mapsto T(\boldsymbol{v}, \boldsymbol{\sigma}_1)$ is continuous. Moreover, for each bounded subset $B \subset \mathbb{R}^{nrM}$, the family of functions $\{T(\boldsymbol{v}; \cdot)\}_{\boldsymbol{v} \in B}$ is equicontinuous.

Proof of Proposition 24. (i) Let $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1$ and $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{nrM}$. For each $(i, s_1, \boldsymbol{p}_0)$-coordinate, let τ_1^i, $\iota_1^i \in \boldsymbol{\Sigma}_1^i$ be such that

$$T(\boldsymbol{u}, \boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} = \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{u})_{i,s_1,\boldsymbol{p}_0} \text{ and}$$
$$T(\boldsymbol{v}, \boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} = \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \iota_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}.$$

From (A.46) and the above equations, it follows that $-T(\boldsymbol{u}, \boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} \leq -\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \iota_1^i, \boldsymbol{u})_{i,s_1,\boldsymbol{p}_0}$ and $-T(\boldsymbol{v}, \boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} \leq -\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}$. Thus,

$$\begin{aligned} &[T(\boldsymbol{u}, \boldsymbol{\sigma}_1) - T(\boldsymbol{v}, \boldsymbol{\sigma}_1)]_{i,s_1,\boldsymbol{p}_0} \leq [\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{u}) - \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \tau_1^i, \boldsymbol{v})]_{i,s_1,\boldsymbol{p}_0} \text{ and} \\ &[T(\boldsymbol{v}, \boldsymbol{\sigma}_1) - T(\boldsymbol{u}, \boldsymbol{\sigma}_1)]_{i,s_1,\boldsymbol{p}_0} \leq [\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \iota_1^i, \boldsymbol{v}) - \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \iota_1^i, \boldsymbol{u})]_{i,s_1,\boldsymbol{p}_0} \end{aligned} \tag{A.47}$$

From (A.47) and (b) in Proposition 23, it follows that

$$\max_{i,s_1,\boldsymbol{p}_0} |T(\boldsymbol{u}, \boldsymbol{\sigma}_1) - T(\boldsymbol{v}, \boldsymbol{\sigma}_1)|_{i,s_1,\boldsymbol{p}_0} \leq \delta \max_{j,s_2,\boldsymbol{p}_1} |v_{j,s_2,\boldsymbol{p}_1} - u_{j,s_2,\boldsymbol{p}_1}|, \tag{A.48}$$

where $\delta = \max_{i \in [n]} \delta_i < 1$. Thus, $T(\cdot, \boldsymbol{\sigma}_1)$ is a contraction mapping. The fact that $T(\cdot, \boldsymbol{\sigma}_1)$ has a unique fixed point follows from Banach Fixed point Theorem.

(ii) Let $\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1 \in \boldsymbol{\Sigma}_1$ and $\boldsymbol{v} \in \mathbb{R}^{nrM}$. For each $(i, s_1, \boldsymbol{p}_0)$-coordinate, let γ_1^i, $\iota_1^i \in \boldsymbol{\Sigma}_1^i$ be such that

$$T(\boldsymbol{v},\boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} = \boldsymbol{V}_1(\boldsymbol{\sigma}_1,\gamma_1^i,\boldsymbol{v})_{i,s_1,\boldsymbol{p}_0} \text{ and}$$
$$T(\boldsymbol{v},\boldsymbol{\tau}_1)_{i,s_1,\boldsymbol{p}_0} = \boldsymbol{V}_1(\boldsymbol{\tau}_1,\iota_1^i,\boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}.$$

From (A.46) and the above equations, it follows that $-T(\boldsymbol{v},\boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} \leq -\boldsymbol{V}_1(\boldsymbol{\sigma}_1,\iota_1^i,\boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}$ and $-T(\boldsymbol{v},\boldsymbol{\tau}_1)_{i,s_1,\boldsymbol{p}_0} \leq -\boldsymbol{V}_1(\boldsymbol{\tau}_1,\gamma_1^i,\boldsymbol{v})_{i,s_1,\boldsymbol{p}_0}$. Thus,

$$\begin{aligned}
&[T(\boldsymbol{v},\boldsymbol{\sigma}_1) - T(\boldsymbol{v},\boldsymbol{\tau}_1)]_{i,s_1,\boldsymbol{p}_0} \leq [\boldsymbol{V}_1(\boldsymbol{\sigma}_1,\gamma_1^i,\boldsymbol{v}) - \boldsymbol{V}_1(\boldsymbol{\tau}_1,\gamma_1^i,\boldsymbol{v})]_{i,s_1,\boldsymbol{p}_0} \text{ and}\\
&[T(\boldsymbol{v},\boldsymbol{\tau}_1) - T(\boldsymbol{v},\boldsymbol{\sigma}_1)]_{i,s_1,\boldsymbol{p}_0} \leq [\boldsymbol{V}_1(\boldsymbol{\tau}_1,\iota_1^i,\boldsymbol{v}) - \boldsymbol{V}_1(\boldsymbol{\sigma}_1,\iota_1^i,\boldsymbol{v})]_{i,s_1,\boldsymbol{p}_0}.
\end{aligned} \tag{A.49}$$

Let $\epsilon > 0$, by part (a) in Proposition 23, there exists $\theta > 0$ such that for each $\kappa \in \{\sigma_1^i, \tau_1^i\}$ if

$$|\boldsymbol{\sigma}_1 - \boldsymbol{\tau}_1|_\infty < \theta \implies |\boldsymbol{V}_1(\boldsymbol{\sigma}_1,\kappa,\boldsymbol{v}) - \boldsymbol{V}_1(\boldsymbol{\tau}_1,\kappa,\boldsymbol{v})|_\infty < \epsilon, \tag{A.50}$$

where $|\boldsymbol{\sigma}_1|_\infty$ denotes the supremum norm of $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1 \subset \mathbb{R}^{n\hat{M}}$ (see (2.4)). From (A.49) and (A.50), it follows that the mapping $\boldsymbol{\sigma}_1 \mapsto T(\boldsymbol{v},\boldsymbol{\sigma}_1)$ is continuous.

Let B be a bounded subset of $\mathbb{R}^{nrM}$. By (2.4), the set $\boldsymbol{\Sigma}_1 \times \boldsymbol{\Sigma}_1 \times \bar{B}$ is compact. By Proposition 23, $\boldsymbol{V}_1$ is uniformly continuous on $\boldsymbol{\Sigma}_1 \times \boldsymbol{\Sigma}_1 \times \bar{B}$. It follows that for each $\epsilon > 0$, there exists $\theta > 0$ such that for each $\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1$ and $\boldsymbol{\kappa}_1$ in $\boldsymbol{\Sigma}_1$ and $\boldsymbol{v} \in B$, if

$$|\boldsymbol{\sigma}_1 - \boldsymbol{\tau}_1|_\infty < \theta \implies |\boldsymbol{V}_1(\boldsymbol{\sigma}_1,\boldsymbol{\kappa}_1,\boldsymbol{v}) - \boldsymbol{V}_1(\boldsymbol{\tau}_1,\boldsymbol{\kappa}_1,\boldsymbol{v})|_\infty < \epsilon. \tag{A.51}$$

Replacing (A.50) with (A.51) shows that the family of functions $\{T(\boldsymbol{v},\cdot)\}_{\boldsymbol{v}\in B}$ is equicontinuous.

□

The mapping b and the correspondence Γ: Let $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1$. From Part (i) in Proposition 24, there exists a unique vector $b(\boldsymbol{\sigma}_1) \in \mathbb{R}^{nrM}$ such that $b(\boldsymbol{\sigma}_1) = T(b(\boldsymbol{\sigma}_1),\boldsymbol{\sigma}_1)$. Thus, there is a well-defined mapping $b : \boldsymbol{\Sigma}_1 \longrightarrow \mathbb{R}^{nrM}$ such that $\boldsymbol{\sigma}_1 \mapsto b(\boldsymbol{\sigma}_1)$. In particular, by (A.46), for each $(i, s_1, \boldsymbol{p}_0)$-coordinate

$$b(\boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} = \max_{\tau_1^i \in \boldsymbol{\Sigma}_1^i} \boldsymbol{V}_1(\boldsymbol{\sigma}_1,\tau_1^i,b(\boldsymbol{\sigma}_1))_{i,s_1,\boldsymbol{p}_0}. \tag{A.52}$$

From (A.52) and the compactness of $\boldsymbol{\Sigma}_1^i$, there exists $\tilde{\boldsymbol{\tau}}_1 \in \boldsymbol{\Sigma}_1$ such that for each $(i, s_1, \boldsymbol{p}_0)$-

coordinate, $b(\boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} = V_1(\boldsymbol{\sigma}_1, \tilde{\tau}_1^i, b(\boldsymbol{\sigma}_1))_{i,s_1,\boldsymbol{p}_0}$. The previous argument shows that for each $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1$, the following set is nonempty,

$$\Gamma(\boldsymbol{\sigma}_1) := \{\boldsymbol{\tau}_1 \in \boldsymbol{\Sigma}_1 | b(\boldsymbol{\sigma}_1) = V_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, b(\boldsymbol{\sigma}_1))\}. \tag{A.53}$$

The mapping Γ from $\boldsymbol{\Sigma}_1$ to $2^{\boldsymbol{\Sigma}_1}$ is a self-correspondence on $\boldsymbol{\Sigma}_1$.

Proposition 25 (Properties of b and Γ)**.** *(i) b is continuous;*

(ii) Γ is a convex- and closed-valued self-correspondence on $\boldsymbol{\Sigma}_1$. Moreover, it has a closed graph.

Proof of Proposition 25. (i) We first show that $b(\boldsymbol{\Sigma}_1) \subset \mathbb{R}^{nrM}$ is bounded. Let $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1$. By Proposition 24, the definition of b and the Banach Fixed Point Theorem: the sequence given by $\boldsymbol{v}_0 = 0 \in \mathbb{R}^{nrM}$ and $\boldsymbol{v}_n = T(\boldsymbol{v}_{n-1}, \boldsymbol{\sigma}_1)$ for $n \geq 1$ converges to $b(\boldsymbol{\sigma}_1)$. Moreover,

$$\max_{i,s_1,\boldsymbol{p}_0} |b(\boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} - 0| \leq \frac{1}{1-\delta} \max_{i,s_1,\boldsymbol{p}_0} |(V_1)_{i,s_1,\boldsymbol{p}_0} - 0|. \tag{A.54}$$

Note that $(V_1)_{i,s_1,\boldsymbol{p}_0} = T(0;\boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0} = \max_{\tau_1^i \in \boldsymbol{\Sigma}_1^i} V_1(\boldsymbol{\sigma}_1, \tau_1^i, 0)_{i,s_1,\boldsymbol{p}_0}$ and by (2.10),

$$V_1(\boldsymbol{\sigma}_1, \tau_1^i, 0)_{i,s_1,\boldsymbol{p}_0} = \sum_{\boldsymbol{p}_1 \in \mathcal{A}^n} \tau_1^i(p_1^i|\boldsymbol{p}_0, s_1)\sigma_1^{-i}(\boldsymbol{p}_1^{-i}|\boldsymbol{p}_0, s_1)\pi^i(\boldsymbol{p}_1, s_1) \tag{A.55}$$

Combining (A.54) with (A.55) yields,

$$\max_{i,s_1,\boldsymbol{p}_0} |b(\boldsymbol{\sigma}_1)_{i,s_1,\boldsymbol{p}_0}| \leq \frac{1}{1-\delta} \max_{i,s_1,\boldsymbol{p}_1} |\pi^i(\boldsymbol{p}_1, s_1)|.$$

Proving that the set $b(\boldsymbol{\Sigma}_1)$ is bounded in $\mathbb{R}^{nrM}$. We now show that b is continuous. Let $\boldsymbol{\sigma}_1$ and $\boldsymbol{\iota}_1$ in $\boldsymbol{\Sigma}_1$. We estimate the following supremum norm

$$\begin{aligned} &|b(\boldsymbol{\sigma}_1) - b(\boldsymbol{\iota}_1)|_\infty = |T(b(\boldsymbol{\sigma}_1), \boldsymbol{\sigma}_1) - T(b(\boldsymbol{\iota}_1), \boldsymbol{\iota}_1)|_\infty \\ &\leq |T(b(\boldsymbol{\sigma}_1), \boldsymbol{\sigma}_1) - T(b(\boldsymbol{\iota}_1), \boldsymbol{\sigma}_1)|_\infty + |T(b(\boldsymbol{\iota}_1), \boldsymbol{\sigma}_1) - T(b(\boldsymbol{\iota}_1), \boldsymbol{\iota}_1)|_\infty. \end{aligned} \tag{A.56}$$

From (A.48) in the Proof of Proposition 24, $|T(b(\boldsymbol{\sigma}_1), \boldsymbol{\sigma}_1) - T(b(\boldsymbol{\iota}_1), \boldsymbol{\sigma}_1)|_\infty \leq \delta |b(\boldsymbol{\sigma}_1) - b(\boldsymbol{\iota}_1)|_\infty$, where $\delta = \max_{i \in [n]} \delta_i$. Thus,

$$|b(\boldsymbol{\sigma}_1) - b(\boldsymbol{\iota}_1)|_\infty \leq \frac{1}{1-\delta} |T(b(\boldsymbol{\iota}_1), \boldsymbol{\sigma}_1) - T(b(\boldsymbol{\iota}_1), \boldsymbol{\iota}_1)|_\infty. \tag{A.57}$$

Since $b(\boldsymbol{\Sigma}_1)$ is bounded, by part (ii) in Proposition 24, the following family of functions, $\{T(\boldsymbol{v};\cdot)\}_{\boldsymbol{v}\in b(\boldsymbol{\Sigma}_1)}$ is equicontinuous. It follows that for each $\epsilon > 0$ there exists $\theta > 0$ such that for any $\boldsymbol{\sigma}_1, \boldsymbol{\iota}_1 \in \boldsymbol{\Sigma}_1$ and $\boldsymbol{v} \in b(\boldsymbol{\Sigma}_1)$, if $|\boldsymbol{\sigma}_1 - \boldsymbol{\iota}_1|_\infty < \theta$, then

$$|T(\boldsymbol{v}, \boldsymbol{\sigma}_1) - T(\boldsymbol{v}, \boldsymbol{\iota}_1)|_\infty < \epsilon(1-\delta).$$

It follows from (A.57) that b is continuous.

(ii) Let $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1$. That $\Gamma(\boldsymbol{\sigma}_1)$ is convex follows from (c) in Proposition 23, as for any $\boldsymbol{\tau}_1, \boldsymbol{\iota}_1 \in \Gamma(\boldsymbol{\sigma}_1)$ and $\alpha \in [0,1]$,

$$\begin{aligned} b(\boldsymbol{\sigma}_1) &= \alpha b(\boldsymbol{\sigma}_1) + (1-\alpha) b(\boldsymbol{\sigma}_1) = \alpha \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, b(\boldsymbol{\sigma}_1)) + (1-\alpha)\boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\iota}_1, b(\boldsymbol{\sigma}_1)) \\ &= \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \alpha\boldsymbol{\tau}_1 + (1-\alpha)\boldsymbol{\iota}_1, b(\boldsymbol{\sigma}_1)). \end{aligned}$$

We now show that $\Gamma(\boldsymbol{\sigma}_1)$ is closed in $\boldsymbol{\Sigma}_1$: Let $(\boldsymbol{\tau}_{1,k})_{k\geq 1} \subset \Gamma(\boldsymbol{\sigma}_1)$ be a sequence such that $\boldsymbol{\tau}_{1,k} \to \boldsymbol{\tau}_1 \in \boldsymbol{\Sigma}_1$ as $k \to \infty$. By definition of $\Gamma(\boldsymbol{\sigma}_1)$ and continuity of $\boldsymbol{V}_1$,

$$b(\boldsymbol{\sigma}_1) = \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_{1,k}, b(\boldsymbol{\sigma}_1)) \to \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\tau}_1, b(\boldsymbol{\sigma}_1)), \text{ as } k \to \infty.$$

It follows that $\boldsymbol{\tau}_1 \in \Gamma(\boldsymbol{\sigma}_1)$.

We show that Γ has a closed graph. Let $(\boldsymbol{\sigma}_{1,k})_k$ and $(\boldsymbol{\iota}_{1,k})_k$ be two sequences in $\boldsymbol{\Sigma}_1$ such that $\boldsymbol{\sigma}_{1,k} \to \boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}_1$ and $\boldsymbol{\iota}_{1,k} \to \boldsymbol{\iota}_1 \in \boldsymbol{\Sigma}_1$ as $k \to \infty$. Suppose that $\boldsymbol{\iota}_{1,k} \in \Gamma(\boldsymbol{\sigma}_{1,k})$ for each $k \geq 1$. By definition, for each $k \geq 1$,

$$b(\boldsymbol{\sigma}_{1,k}) = \boldsymbol{V}_1(\boldsymbol{\sigma}_{1,k}, \boldsymbol{\iota}_{1,k}, b(\boldsymbol{\sigma}_{1,k})).$$

By part (i) in this proposition, b is continuous, therefore $b(\boldsymbol{\sigma}_{1,k}) \to b(\boldsymbol{\sigma}_1)$ as $k \to \infty$. By Proposition 23, $\boldsymbol{V}_1$ is continuous, thus, $\boldsymbol{V}_1(\boldsymbol{\sigma}_{1,k}, \boldsymbol{\iota}_{1,k}, b(\boldsymbol{\sigma}_{1,k})) \to \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\iota}_1, b(\boldsymbol{\sigma}_1))$ as $k \to \infty$. It follows that

$$b(\boldsymbol{\sigma}_1) = \boldsymbol{V}_1(\boldsymbol{\sigma}_1, \boldsymbol{\iota}_1, b(\boldsymbol{\sigma}_1)),$$

and $\boldsymbol{\iota}_1 \in \Gamma(\boldsymbol{\sigma}_1)$.

□

Conclusion of Theorem 1

By Proposition 25, Γ as given by (A.53) is a convex-valued self-correspondence on $\mathbf{\Sigma}_1$ that has a closed graph. Moreover, $\mathbf{\Sigma}_1$ is compact and convex. By Theorem 5, there exists $\boldsymbol{\sigma}_1^* \in \mathbf{\Sigma}_1$ such that $\boldsymbol{\sigma}_1^* \in \Gamma(\boldsymbol{\sigma}_1^*)$, i.e.,

$$b(\boldsymbol{\sigma}_1^*) = \boldsymbol{V}_1(\boldsymbol{\sigma}_1^*, \boldsymbol{\sigma}_1^*, b(\boldsymbol{\sigma}_1^*)).$$

Appendix B

Proofs of Chapter 3

We prove all the stated results in the following order: Proposition 8, Proposition 9 (for which we first provide various definitions and establish Lemma 2), Proposition 10 (for which we first prove Lemma 3), Proposition 11 (for which we first prove Lemma 4), Proposition 12 (for which we first prove Lemma 5), Proposition 13, Corollary 1, Corollary 2, Proposition 14, Proposition 15, Proposition 16, Proposition 17, Corollary 3, Corollary 4, Proposition 18, Proposition 19, Proposition 20, Proposition 21 and Proposition 22. In order to save space, we leave some of the lengthy calculations to Mathematica and report them in the supplementary file *Gumbel_N.nb*.

Proof of Proposition 8. Let $\{(p_b^i, p_s^i)\}_{i=1}^N \subset \mathbb{R}^2$ be a set of prices. For $i \in \mathcal{N} \cup \{0\}$ and $k \in \{b, s\}$, set $v_k^i := \hat{u}_k^i - \varepsilon_k^i$. From (3.1) and (3.2), for $i, j \in \mathcal{N} \cup \{0\}$, $i \neq j$, (3.3) can be rewritten as

$$\begin{aligned} x_k^i &= \mathbb{P}(\hat{u}_k^i > \max_{j \neq i}(\hat{u}_k^j)) \\ &= \mathbb{P}(\epsilon_k^i > \max_{j \neq i}(\epsilon_k^j + v_k^j - v_k^i)), \quad k \in \{b, s\}. \end{aligned} \tag{B.1}$$

For $i \in \mathcal{N} \cup \{0\}$ and $k \in \{b, s\}$, we define $T_k^i : \mathbb{R}^{N+1} \longrightarrow [0, 1]$ such that

$$\boldsymbol{u} = (u^0, u^1, \cdots, u^N) \mapsto T_k^i(\boldsymbol{u}) := \mathbb{P}(\underbrace{\epsilon_k^i > \max_{j \neq i}(\epsilon_k^j + u^j - u^i)}_{:= E_k^i(\boldsymbol{u})}).$$

Note that $E_k^i(\boldsymbol{u}) \subset \Omega$ (where Ω is the domain of the random variables $\{\varepsilon_k^i\}_{i \in \mathcal{N} \cup \{0\}, k \in \{b,s\}}$). In two steps, we show that for any $\boldsymbol{u} \in \mathbb{R}^{N+1}$ and $k \in \{b, s\}$, $\sum_{i=0}^N T_k^i(\boldsymbol{u}) = 1$.

Step (i): For any $i \neq j$, the events $E_k^i(\boldsymbol{u})$ and $E_k^j(\boldsymbol{u})$ are disjoint because either $\epsilon_k^i > \epsilon_k^j + u^j - u^i$ or $\epsilon_k^j > \epsilon_k^i + u^i - u^j$, but not both. Then, $\sum_{i=0}^N T_k^i(\boldsymbol{u}) = \mathbb{P}\left(\cup_{i=0}^N E_k^i\right)$.

Step (ii): We show that $\mathbb{P}\left(\cap_{i=0}^N (E_k^i)^c\right) = 0$. First, note that the following sets $\{\overline{E_k^i} \cap (E_k^i)^c\}_{i \in \mathcal{N} \cup \{0\}, k \in \{b,s\}}$ have $\mathbb{P}$-zero probability, because $\mathbb{P}$ is absolutely continuous with respect to the Lebesgue measure and each of the sets $\overline{E_k^i} \cap (E_k^i)^c$ is contained inside an N-dimensional set of $\mathbb{R}^{N+1}$. We claim that $\cap_{i=0}^N (E_k^i)^c \subseteq \cup_{i=0}^N (\overline{E_k^i} \cap (E_k^i)^c)$. Let $\omega \in \cap_{i=0}^N (E_k^i)^c$. Then, for all $i \in \mathcal{N} \cup \{0\}$,

$$\epsilon_k^i \leq \max_{j \neq i} (\epsilon_k^j + u^j - u^i). \tag{B.2}$$

If there exists $i \in \mathcal{N} \cup \{0\}$ such that (B.2) holds with equality, then $\omega \in \overline{E_k^i} \cap (E_k^i)^c$ and the claim holds true. Now we prove that if for all $i \in \mathcal{N} \cup \{0\}$, (B.2) is satisfied with strict inequality, we get a contradiction. By (B.2) with strict inequality, there exists $\theta(0) \in \mathcal{N} \cup \{0\}$, $\theta(0) \neq 0$ such that $\epsilon_k^0 < \epsilon_k^{\theta(0)} + u^{\theta(0)} - u^0$. Similarly, there exists $\theta^2(0) \neq \theta(0)$ such that $\epsilon_k^{\theta(0)} < \epsilon_k^{\theta^2(0)} + u^{\theta^2(0)} - u^{\theta(0)}$. Note that $\theta^2(0) \neq 0$, otherwise, $\epsilon_k^{\theta(0)} < \epsilon_k^0 + u^0 - u^{\theta(0)}$ which contradicts the definition of $\theta(0)$. By induction, suppose that for $n \in \mathbb{N}$ and all $0 \leq m \leq n$, there exists $\theta^m(0) \in \mathcal{N} \cup \{0\}$ such that

$$\theta^m(0) \notin \{0, \theta(0), \cdots, \theta^{m-1}(0)\} \text{ and } \epsilon_k^{\theta^{m-1}(0)} < \epsilon_k^{\theta^m(0)} + u^{\theta^m(0)} - u^{\theta^{m-1}(0)}. \tag{B.3}$$

By (B.2) with strict inequality, there exists $\theta^{n+1}(0) \in \mathcal{N} \cup \{0\}$, $\theta^{n+1}(0) \neq \theta^n(0)$ such that $\epsilon_k^{\theta^n(0)} < \epsilon_k^{\theta^{n+1}(0)} + u^{\theta^{n+1}(0)} - u^{\theta^n(0)}$. We claim that $\theta^{n+1}(0) \notin \{0, \theta(0), \cdots, \theta^n(0)\}$, otherwise $\theta^{n+1}(0) = \theta^m(0)$ for some $0 \leq m \leq n-1$. In this case, by (B.3)

$$\begin{aligned} \epsilon_k^{\theta^{n+1}(0)} = \epsilon_k^{\theta^m(0)} &< \epsilon_k^{\theta^{m+1}(0)} + u^{\theta^{m+1}(0)} - u^{\theta^m(0)} \\ &< \epsilon_k^{\theta^{m+2}(0)} + u^{\theta^{m+2}(0)} - \cancel{u^{\theta^{m+1}(0)}} + \cancel{u^{\theta^{m+1}(0)}} - u^{\theta^m(0)} \\ &\vdots \\ &< \epsilon_k^{\theta^n(0)} + u^{\theta^n(0)} - u^{\theta^m(0)}. \end{aligned} \tag{B.4}$$

Note that (B.4) contradicts the definition of $\theta^{n+1}(0)$. Then $\theta^{n+1}(0) \notin \{0, \theta(0), \cdots, \theta^n(0)\}$ and (B.3) is satisfied for the next index $n+1$. It follows that (B.3) holds for any $n \in \mathbb{N}$. The latter is impossible because there are only $N+1$ different indices inside $\mathcal{N} \cup \{0\}$. Thus, $\cap_{i=0}^N (E_k^i)^c \subseteq \cup_{i=0}^N (\overline{E_k^i} \cap (E_k^i)^c)$ and $\mathbb{P}\left(\cap_{i=0}^N (E_k^i)^c\right) = 0$.

Combining steps (i) and (ii), we get that for any $\boldsymbol{u} \in \mathbb{R}^{N+1}$ and $k \in \{b, s\}$, the following equality holds true, $\sum_{i=0}^{2} T_k^i(\boldsymbol{u}) = 1$. Now, for $\boldsymbol{x} = (x_{0b}, x_{0s}, x_{1b}, x_{1s}, \ldots, x_{Nb}, x_{Ns}) \in [0,1]^{2(N+1)}$ and each $i \in \mathcal{N} \cup \{0\}$, we introduce the auxiliary functions $\phi_k^i(\boldsymbol{x})$ defined as

$$\phi_k^i(\boldsymbol{x}) = \begin{cases} \phi_k^0 & \text{if } i = 0 \\ \phi_k(x_{ib}, x_{is}) & \text{if } i \geq 1 \end{cases}.$$

Similarly, we define

$$\sigma_k^i(\boldsymbol{x}) := T_k^i(v_k^0(\boldsymbol{x}), \cdots, v_k^N(\boldsymbol{x})),$$

where $v_k^j(\boldsymbol{x}) = \phi_k^j(\boldsymbol{x}) - p_k^j$ ($p_k^0 = 0$). If $\Sigma : [0,1]^{2(N+1)} \longrightarrow [0,1]^{2(N+1)}$ is defined by

$$\Sigma(\boldsymbol{x}) = (\sigma_b^0(\boldsymbol{x}), \sigma_s^0(\boldsymbol{x}) \cdots, \sigma_b^N(\boldsymbol{x}), \sigma_s^N(\boldsymbol{x})),$$

then solving system (3.3) is equivalent to finding a fixed point of Σ, i.e., $\Sigma(\boldsymbol{x}) = \boldsymbol{x}$. Existence of such a fixed-point is guaranteed by Brouwer's Fixed Point Theorem, as Σ is continuous on $[0,1]^{2(N+1)}$. For such a fixed point:

$$\sum_{i=0}^{N} x_k^i = \sum_{i=0}^{N} \sigma_k^i(\boldsymbol{x}) = \sum_{i=0}^{N} T_k^i(v_k^0(\boldsymbol{x}), \cdots, v_k^N(\boldsymbol{x})) = 1$$

.

To show the uniqueness of the solution of (3.3), we use the Banach Fixed Point Theorem. Let $\boldsymbol{x}, \boldsymbol{y} \in [0,1]^{2(N+1)}$, then

$$\begin{aligned} |\sigma_k^i(\boldsymbol{x}) - \sigma_k^i(\boldsymbol{y})| &= |T_k^i(v_k^0(\boldsymbol{x}), \cdots, v_k^N(\boldsymbol{x})) - T_k^i(v_k^0(\boldsymbol{y}), \cdots, v_k^N(\boldsymbol{y}))| \\ &\leq \max_j |v_k^j(\boldsymbol{x}) - v_k^j(\boldsymbol{y})| \cdot \left(\sup_{\boldsymbol{u} \in \mathbb{R}^{N+1}} \sum_{l=0}^{N} \left| \frac{\partial T_k^i(\boldsymbol{u})}{\partial u^l} \right| \right) \\ &\leq \max_j |\phi_k^j(\boldsymbol{x}) - \phi_k^j(\boldsymbol{y})| \cdot M_T \\ &\leq M_\phi M_T |\boldsymbol{x} - \boldsymbol{y}|_\infty, \end{aligned}$$

where

$$M_T := \max_{k\in\{b,s\},i\in\mathcal{N}\cup\{0\}} \sup_{\boldsymbol{u}\in\mathbb{R}^{N+1}} \sum_{l=0}^{N} \left|\frac{\partial T_k^i(\boldsymbol{u})}{\partial u^l}\right|, \text{ and}$$
$$M_\phi := \max_{k\in\{b,s\}} \sup_{(x_b,x_s)\in[0,1]^2} \sum_{l\in\{b,s\}} \left|\frac{\partial \phi_k(x_b,x_s)}{\partial x_l}\right|.$$

It follows that $\Sigma(\cdot)$ is a (strict) contracting mapping whenever $M_T M_\phi < 1$, and uniqueness follows.

□

Preliminary results for the proof of Proposition 9. We introduce notation and definitions and establish a useful lemma. Let $\boldsymbol{X} = \left(x_b^1, \cdots, x_b^N, x_s^1, \cdots, x_s^N\right)$ and $\boldsymbol{P} = \left(p_b^1, \cdots, p_b^N, p_s^1, \cdots, p_s^N\right)$ be two vectors in $\mathbb{R}^{2N}$. For $k \in \{b, s\}$, let $\tilde{\boldsymbol{u}}_k := (u_k^0, u_k^1, \ldots, u_k^N)$, where $u_k^i = \phi_k\left(x_b^i, x_s^i\right) - p_k^i$. Using (3.9), we can define a mapping from $\mathbb{R}^{4N}$ to $\mathbb{R}^{2N}$ as

$$\begin{aligned}&(\boldsymbol{X}, \boldsymbol{P}) \mapsto \mathcal{T}(\boldsymbol{X}, \boldsymbol{P}) := \\ &\left(T_b^1(\tilde{\boldsymbol{u}}_b) - x_b^1, \cdots, T_b^N(\tilde{\boldsymbol{u}}_b) - x_b^N, T_s^1(\tilde{\boldsymbol{u}}_s) - x_s^1, \cdots, T_s^N(\tilde{\boldsymbol{u}}_s) - x_s^N\right).\end{aligned} \tag{B.5}$$

The Jacobian of (3.9) w.r.t. $\boldsymbol{P}$ is defined as

$$\begin{aligned}\det\frac{\partial\mathcal{T}}{\partial\boldsymbol{P}}(\boldsymbol{X}, \boldsymbol{P}) &:= Q_b(\boldsymbol{X}, \boldsymbol{P})\, Q_s(\boldsymbol{X}, \boldsymbol{P}), \quad \text{where} \\ Q_k(\boldsymbol{X}, \boldsymbol{P}) &:= \begin{vmatrix} -\frac{\partial T_k^1}{\partial u^1}(\tilde{\boldsymbol{u}}_k) & \ldots & -\frac{\partial T_k^1}{\partial u^N}(\tilde{\boldsymbol{u}}_k) \\ \vdots & \ddots & \vdots \\ -\frac{\partial T_k^N}{\partial u^1}(\tilde{\boldsymbol{u}}_k) & \ldots & -\frac{\partial T_k^N}{\partial u^N}(\tilde{\boldsymbol{u}}_k) \end{vmatrix}.\end{aligned} \tag{B.6}$$

Under symmetry, for any $i \in \mathcal{N}$ and $k \in \{b, s\}$, we write $p_k^i = p_k$, $x_k^i = x_k$, and $u_k^i = u_k := \phi_k(x_b, x_s) - p_k$. Let $\boldsymbol{u}_k := (u_k^0, \phi_k(x_b, x_s) - p_k, \cdots, \phi_k(x_b, x_s) - p_k)^T \in \mathbb{R}^{N+1}$.

For $i, j \in \mathcal{N}$, $i \neq j$, $k \in \{b, s\}$, we define the functions

$$
\begin{aligned}
&S_k(\boldsymbol{u}_k) := \frac{\partial T_k^i}{\partial u_k^i}(\boldsymbol{u}_k), \\
&R_k(\boldsymbol{u}_k) := \frac{\partial T_k^i}{\partial u_k^j}(\boldsymbol{u}_k), \\
&J_k(\boldsymbol{u}_k) := S_k(\boldsymbol{u}_k)\left(S_k(\boldsymbol{u}_k) + (N-2)R_k(\boldsymbol{u}_k)\right) - (N-1)R_k(\boldsymbol{u}_k)^2, \text{ and} \\
&J_\phi(\boldsymbol{u}_b, \boldsymbol{u}_s) := \left(\frac{\partial \phi_s}{\partial x_s} - \frac{1}{J_s(\boldsymbol{u}_s)}S_s(\boldsymbol{u}_s)\right)\left(\frac{\partial \phi_b}{\partial x_b} - \frac{1}{J_b(\boldsymbol{u}_b)}S_b(\boldsymbol{u}_b)\right) - \frac{\partial \phi_s}{\partial x_b}\frac{\partial \phi_b}{\partial x_s}.
\end{aligned}
\tag{B.7}
$$

Whenever there is no room for confusion, we simplify the notations by neglecting the explicit mention of the input $\boldsymbol{u}_k$. For example, $T_k^i(\boldsymbol{u}_k)$, $S_k(\boldsymbol{u}_k)$, $R_k(\boldsymbol{u}_k)$, $J_k(\boldsymbol{u}_k)$ and $J_\phi(\boldsymbol{u}_b, \boldsymbol{u}_s)$ are simplified to T_k^i, S_k, R_k, J_k and J_ϕ respectively. The following Lemma shows the first-order condition of (3.4) as a function of x_k.

Lemma 2 (FOC of CNE)**.** *If* $\det\frac{\partial \mathcal{T}}{\partial \boldsymbol{P}}(\boldsymbol{x}^*, \boldsymbol{p}^*) \neq 0$*, then the symmetric Nash equilibrium outputs* $\mathbf{p}^*$ *and* $\boldsymbol{x}^*$ *are solutions of (3.3) and of the following two equations*

$$
\begin{aligned}
&p_k + \frac{\partial \phi_k}{\partial x_k}x_k + \frac{\partial \phi_l}{\partial x_k}x_l - \frac{1}{J_k}\left(S_k + (N-2)R_k\right)x_k + \frac{1}{J_k^2 J_l J_\phi}(N-1)R_k^2 S_l x_k \\
&+ \frac{N-1}{J_k J_\phi}R_k\left(\frac{1}{J_l}R_l\frac{\partial \phi_l}{\partial x_k}x_l - \frac{1}{J_k}R_k\frac{\partial \phi_l}{\partial x_l}x_k\right) = 0, \text{ for } k, l \in \{b, s\}, \ k \neq l.
\end{aligned}
\tag{B.8}
$$

The proof of this Lemma does not require assumptions I and II of Section 3.2. Thus, the FOC given by (B.8) is applicable to idiosyncratic preferences other than Gumbel distribution and to more general externality functions $\phi_k(\boldsymbol{x})$.

Proof of Lemma 2. Assume that all platforms follow a symmetric equilibrium where $\boldsymbol{p}^i = \boldsymbol{p}^* = (p_b^*, p_s^*)$ and $\boldsymbol{x}^i = \boldsymbol{x}^* = (x_b^*, x_s^*)$. We show that unilateral deviations from this strategy lead to zero gain. Without loss of generality, we assume that the first platform deviates from the symmetric setting. This platform can deviate by either choosing prices $p_k^1 \neq p_k^*$ or market shares $x_k^1 \neq x_k^*$. Suppose that $\det\frac{\partial \mathcal{T}}{\partial \boldsymbol{P}}(\boldsymbol{X}^*, \boldsymbol{P}^*) \neq 0$, where

$$
\boldsymbol{X}^* = (x_b^*, \cdots, x_b^*, x_s^*, \cdots, x_s^*) \text{ and } \boldsymbol{P}^* = (p_b^*, \cdots, p_b^*, p_s^*, \cdots, p_s^*)
$$

belong to $\mathbb{R}^{2N}$. Then, by the Implicit Function Theorem, there exists a neighborhood B of $\boldsymbol{X}^*$ in $\mathbb{R}^{2N}$ and a unique differentiable function $\mathcal{P} : B \longrightarrow \mathbb{R}^{2N}$ such that $\mathcal{P}(\boldsymbol{X}^*) = \boldsymbol{P}^*$

and

$$\mathcal{T}\left(\boldsymbol{X}, \mathcal{P}\left(\boldsymbol{X}\right)\right) = 0 \text{ for all } \boldsymbol{X} \in B. \tag{B.9}$$

From (3.4) and (B.9), we can compute the FOC for this platform w.r.t. x_k^1 as

$$\left.\frac{\partial \pi^1}{\partial x_k^1}\right|_{\boldsymbol{p}^i=\boldsymbol{p}^*, \boldsymbol{x}^i=\boldsymbol{x}^*, \text{ for } i\neq 1} = p_k^1 + x_k^1 \frac{\partial p_k^1}{\partial x_k^1} + x_l^1 \frac{\partial p_l^1}{\partial x_k^1} = 0, \quad \text{for each } k, l \in \{b, s\}, \; k \neq l. \tag{B.10}$$

To solve (B.10), we need to compute the following derivatives

$$\frac{\partial p_k^1}{\partial x_l^1}, \quad \text{for each } k, l \in \{b, s\}. \tag{B.11}$$

We determine those four partial derivatives $\frac{\partial \boldsymbol{p}^1}{\partial \boldsymbol{x}^1}$ in (B.11) using the definition of T_k in (3.8). By (3.9), for $k \in \{b, s\}$, the vectors of market shares and prices, $(x_k^1, x_k^*, \ldots, x_k^*)$ and $(p_k^1, p_k^*, \ldots, p_k^*)$ satisfy all the following

$$T_k^1(u_k^0, u_k^1, u_k^2, \ldots, u_k^N) = x_k^1 \quad \text{and} \tag{B.12}$$

$$T_k^i(u_k^0, u_k^1, u_k^2, \ldots, u_k^N) = x_k^i, \quad \text{for } i \in \{2, 3, \cdots, N\}. \tag{B.13}$$

Taking derivatives w.r.t. x_b^1 and x_s^1 in (B.12) and (B.13), respectively, gives us

$$\frac{\partial T_k^1(u_k^0, u_k^1, u_k^2, \ldots, u_k^N)}{\partial x_l^1} = \delta_{kl} \quad \text{and} \tag{B.14}$$

$$\frac{\partial T_k^i(u_k^0, u_k^1, u_k^2, \ldots, u_k^N)}{\partial x_l^1} = \frac{\partial x_k^i}{\partial x_l^1}, \quad \text{for } i \in \{2, 3, \cdots, N\}, \; k, l \in \{b, s\}, \tag{B.15}$$

where $\delta_{kl} = 1$ if $k = l$ and $\delta_{kl} = 0$ if $k \neq l$. Note that the system of equations in(B.14) and (B.15) includes $4 + 4(N-1) = 4N$ equations. The unknowns are $\frac{\partial p_k^1}{\partial x_l^1}$ and $\frac{\partial x_k^i}{\partial x_l^1}$, for $k, l \in \{b, s\}$ and $i \in \{2, 3, \ldots, N\}$, adding up to $4N$ unknowns.

Note that u_k^0 is a constant and it does not depend on x_k^1. By the chain rule, the left hand side of (B.14) and (B.15) can be rewritten, for $k, l \in \{b, s\}, \; j \in \{1, 2, \ldots, N\}$, as

$$\frac{\partial T_k^j(u_k^0, u_k^1, u_k^2, \ldots, u_k^N)}{\partial u_1} \frac{\partial u_k^1}{\partial x_l^1} + \sum_{i=2}^{N} \frac{\partial T_k^j(u_k^0, u_k^1, u_k^2, \ldots, u_k^N)}{\partial u_i} \frac{\partial u_k^i}{\partial x_l^1}, \tag{B.16}$$

Recall that $u_k^1 = \phi_k(\boldsymbol{x}^1) - p_k^1$, then, we can explicitly write

$$\frac{\partial u_k^1}{\partial x_l^1} = -\frac{\partial p_k^1}{\partial x_l^1} + \frac{\partial \phi_k}{\partial x_l}, \qquad \text{for } k, l \in \{b, s\}. \tag{B.17}$$

On the other hand, for $i \in \{2, 3, \cdots, N\}$, $u_k^i = \phi_k(\boldsymbol{x}^i) - p_k^*$. Thus,

$$\frac{\partial u_k^i}{\partial x_l^1} = \frac{\partial \phi_k}{\partial x_b}\frac{\partial x_b^i}{\partial x_l^1} + \frac{\partial \phi_k}{\partial x_s}\frac{\partial x_s^i}{\partial x_l^1}, \qquad \text{for } k, l \in \{b, s\},\ i \in \{2, \ldots, N\}. \tag{B.18}$$

Plugging (B.16), (B.17) and (B.18) into (B.14) and (B.15), we obtain for $k, l \in \{b, s\}$ and $j \in \{2, 3, \cdots, N\}$,

$$\frac{\partial T_k^1}{\partial u_1}\left(-\frac{\partial p_k^1}{\partial x_l^1} + \frac{\partial \phi_k}{\partial x_l}\right) + \sum_{i=2}^{N}\frac{\partial T_k^1}{\partial u_i}\left(\frac{\partial \phi_k}{\partial x_b}\frac{\partial x_b^i}{\partial x_l^1} + \frac{\partial \phi_k}{\partial x_s}\frac{\partial x_s^i}{\partial x_l^1}\right) = \delta_{kl}, \tag{B.19}$$

$$\frac{\partial T_k^j}{\partial u_1}\left(-\frac{\partial p_k^1}{\partial x_l^1} + \frac{\partial \phi_k}{\partial x_l}\right) + \sum_{i=2}^{N}\frac{\partial T_k^j}{\partial u_i}\left(\frac{\partial \phi_k}{\partial x_b}\frac{\partial x_b^i}{\partial x_l^1} + \frac{\partial \phi_k}{\partial x_s}\frac{\partial x_s^i}{\partial x_l^1}\right) = \frac{\partial x_k^j}{\partial x_l^1}, \tag{B.20}$$

There are $4N$ equations with $4N$ variables in the above system. Before solving the system, we want to apply the property of symmetry at equilibrium, where we denote

$$\begin{aligned} &\frac{\partial x_k^j}{\partial x_l^1} =: y_{k,l}, \quad \text{for } j \in \{2, 3, \cdots, N\} \text{ and} \\ &\frac{\partial T_k^i}{\partial u_j} =: \begin{cases} S_k & i = j \\ R_k & i \neq j \end{cases}, \quad \text{for } i, j \in \{1, \ldots, N\}. \end{aligned} \tag{B.21}$$

Incorporating (B.21), we can further reduce the system described by (B.19) and (B.20) into 8 equations with unknowns: $\frac{\partial p_k^1}{\partial x_l^1}$ and $y_{k,l}$, for $k, l \in \{b, s\}$. The new system is, for $k, l \in \{b, s\}$,

$$S_k\left(-\frac{\partial p_k^1}{\partial x_l^1} + \frac{\partial \phi_k}{\partial x_l}\right) + (N-1)\,R_k\left(\frac{\partial \phi_k}{\partial x_b}y_{b,l} + \frac{\partial \phi_k}{\partial x_s}y_{s,l}\right) = \delta_{kl} \text{ and} \tag{B.22}$$

$$R_k\left(-\frac{\partial p_k^1}{\partial x_l^1} + \frac{\partial \phi_k}{\partial x_l}\right) + (S_k + (N-2)\,R_k)\left(\frac{\partial \phi_k}{\partial x_b}y_{b,l} + \frac{\partial \phi_k}{\partial x_s}y_{s,l}\right) = y_{k,l}. \tag{B.23}$$

To solve the above equations, we notice that there are two groups of equations: (A) four equations associated to derivatives w.r.t. x_b^1; (B) four equations associated to derivatives w.r.t. x_s^1. Thus, we solve two linear systems, each one containing four equations and four unknowns.

Using (B.7), we conclude that when $J_k \neq 0$ for any $k \in \{b, s\}$ and $J_\phi \neq 0$,

$$y_{b,b} = -\frac{1}{J_b J_\phi} R_b \left(\frac{\partial \phi_s}{\partial x_s} - \frac{S_s}{J_s} \right), \tag{B.24}$$

$$y_{s,b} = \frac{1}{J_b J_\phi} R_b \frac{\partial \phi_s}{\partial x_b}, \tag{B.25}$$

$$y_{b,s} = \frac{1}{J_s J_\phi} R_s \frac{\partial \phi_b}{\partial x_s}, \quad \text{and} \tag{B.26}$$

$$y_{s,s} = -\frac{1}{J_s J_\phi} R_s \left(\frac{\partial \phi_b}{\partial x_b} - \frac{S_b}{J_b} \right). \tag{B.27}$$

Moreover, it follows that

$$\frac{\partial p_b^1}{\partial x_b^1} = \frac{\partial \phi_b}{\partial x_b} - \frac{1}{J_b}(S_b + (N-2)R_b) + \frac{N-1}{J_b} R_b \left(-\frac{1}{J_b J_\phi} R_b \left(\frac{\partial \phi_s}{\partial x_s} - \frac{1}{J_s} S_s \right) \right), \tag{B.28}$$

$$\frac{\partial p_s^1}{\partial x_b^1} = \frac{\partial \phi_s}{\partial x_b} + \frac{N-1}{J_s} R_s \left(\frac{1}{J_b J_\phi} R_b \frac{\partial \phi_s}{\partial x_b} \right), \tag{B.29}$$

$$\frac{\partial p_b^1}{\partial x_s^1} = \frac{\partial \phi_b}{\partial x_s} + \frac{N-1}{J_b} R_b \left(\frac{1}{J_s J_\phi} R_s \frac{\partial \phi_b}{\partial x_s} \right), \quad \text{and} \tag{B.30}$$

$$\frac{\partial p_s^1}{\partial x_s^1} = \frac{\partial \phi_s}{\partial x_s} - \frac{1}{J_s}(S_s + (N-2)R_s) + \frac{N-1}{J_s} R_s \left(-\frac{1}{J_s J_\phi} R_s \left(\frac{\partial \phi_b}{\partial x_b} - \frac{1}{J_b} S_b \right) \right). \tag{B.31}$$

Finally, we can plug in (B.28), (B.30), (B.29) and (B.31) into the FOC (B.10) to obtain

$$\begin{aligned} p_k + \frac{\partial \phi_k}{\partial x_k} x_k + \frac{\partial \phi_l}{\partial x_k} x_l - \frac{1}{J_k}(S_k + (N-2)R_k) x_k + \frac{1}{J_k^2 J_l J_\phi}(N-1) R_k^2 S_l x_k \\ + \frac{N-1}{J_k J_\phi} R_k \left(\frac{1}{J_l} R_l \frac{\partial \phi_l}{\partial x_k} x_l - \frac{1}{J_k} R_k \frac{\partial \phi_l}{\partial x_l} x_k \right) = 0, \quad \text{for } k, l \in \{b, s\},\ k \neq l. \end{aligned} \tag{B.32}$$

□

Lemma 2 is general enough to accommodate idiosyncratic preferences other than Gumbel distribution and general externality functions $\phi_k(\boldsymbol{x})$. Next, we use Assumptions I and

II from Section 3.2 and Lemma 2 to prove Proposition 9.

Proof of Proposition 9. We want to rewrite the FOC given by (B.8) using Assumptions I and II from Section 3.2. Applying Gumbel distribution, we can derive specific forms for the functions T_k^i, S_k, and R_k as defined in (B.7). By Assumption I, $\{\varepsilon_k^i\}_{k\in\{b,s\},i\in\mathcal{N}\cup\{0\}}$ are i.i.d. Gumbel distributed with distribution

$$F_k(z) = e^{-e^{\frac{\mu_k - z}{\beta_k}}}. \tag{B.33}$$

For any $i \in \mathcal{N}$, the random variable $Y^i := \epsilon_k^i + u_k^i - \max_{j=0,1,\ldots,N, j\neq i}\left\{\epsilon_k^j + u_k^j\right\}$ has a logistic distribution,

$$Y^i \sim \text{Logistic}(u_k^i - \beta_k \ln \alpha_i, \beta_k), \quad \text{where} \quad \alpha_i := \sum_{j=0,1,\ldots,N, j\neq i} e^{\frac{u_k^j}{\beta_k}}.$$

By (3.8), we can explicitly write T_k^i as follows

$$\begin{aligned} T_k^i(u_k^0, u_k^1, \ldots, u_k^N) &= \mathbb{P}\left(\epsilon_k^i + u_k^i \geq \max_{j=0,1,\ldots N, j\neq i}\left\{\epsilon_k^j + u_k^j\right\}\right) \\ &= 1 - F_{Y^i}(0) = 1 - \frac{1}{1 + e^{\left(u_k^i - \beta_k \ln(\alpha_i)\right)/\beta_k}}. \end{aligned} \tag{B.34}$$

The derivatives of T_k^i can be calculated as

$$\frac{\partial T_k^i(u_k^0, u_k^1, \ldots, u_k^N)}{\partial u^i} = \frac{\frac{1}{\beta_k} e^{\left(u_k^i - \beta_k \ln(\alpha_i)\right)/\beta_k}}{\left(1 + e^{\left(u_k^i - \beta_k \ln(\alpha_i)\right)/\beta_k}\right)^2} \quad \text{and} \tag{B.35}$$

$$\frac{\partial T_k^i(u_k^0, u_k^1, \ldots, u_k^N)}{\partial u^j} = \frac{-e^{\left(u_k^i - \beta_k \ln(\alpha_i)\right)/\beta_k} \cdot \frac{\frac{1}{\beta_k} e^{u_k^j/\beta_k}}{\alpha_i}}{\left(1 + e^{\left(u_k^i - \beta_k \ln(\alpha_i)\right)/\beta_k}\right)^2}, \quad \text{for } j \neq i. \tag{B.36}$$

At a symmetric equilibrium, $\boldsymbol{u}^i = \boldsymbol{u} = (u_b, u_s)^T$, for any $i \in \mathcal{N}$. Then, we can further simplify (B.34), (B.35) and (B.36),

$$T_k^i = \frac{e^{u_k/\beta_k}}{e^{u_k^0/\beta_k} + Ne^{u_k/\beta_k}}, \tag{B.37}$$

$$S_k = \frac{1}{\beta_k}\frac{e^{u_k/\beta_k}(e^{u_k^0/\beta_k} + (N-1)e^{u_k/\beta_k})}{\left(e^{u_k^0/\beta_k} + Ne^{u_k/\beta_k}\right)^2} \text{ and} \tag{B.38}$$

$$R_k = -\frac{1}{\beta_k}\frac{e^{2u_k/\beta_k}}{\left(e^{u_k^0/\beta_k} + Ne^{u_k/\beta_k}\right)^2}. \tag{B.39}$$

Using (3.12), we derive the dependence of $\boldsymbol{x}$ and $\boldsymbol{p}$ on $\boldsymbol{z}$ under the symmetric setting as

$$x_k^i = T_k^i(\boldsymbol{u}) = \frac{e^{u_k/\beta_k}}{e^{u_k^0/\beta_k} + Ne^{u_k/\beta_k}} = \frac{1}{e^{(u_k^0-u_k)/\beta_k} + N} = \frac{1}{e^{-z_k} + N} =: \omega(z_k). \tag{B.40}$$

Denoting $\Omega(\boldsymbol{z}) := (\omega(z_b), \omega(z_s))^T$, we obtain $\boldsymbol{x} = \Omega(\boldsymbol{z})$ in the symmetric equilibrium. Moreover, we can write $\boldsymbol{u} - \boldsymbol{u}_0 = \boldsymbol{\beta z}$, $\boldsymbol{u} = \Phi\boldsymbol{x} - \boldsymbol{p}$ and

$$\boldsymbol{p} = \Phi\Omega(\boldsymbol{z}) - \boldsymbol{\beta z} - \boldsymbol{u}_0. \tag{B.41}$$

At this point, we want to use Lemma 2 to rewrite (B.8) using Assumptions I and II. First, we verify that $\det\frac{\partial\mathcal{T}}{\partial\boldsymbol{P}}(\boldsymbol{X}, \boldsymbol{P}) \neq 0$, see (B.6), for any pair of symmetric vectors

$$\boldsymbol{X} = (x_b, \cdots, x_b, x_s, \cdots, x_s) \text{ and } \boldsymbol{P} = (p_b, \cdots, p_b, p_s, \cdots, p_s).$$

Using (B.35), (B.36) and (3.12) into (B.6), we obtain

$$\begin{aligned}
Q_k(\boldsymbol{X}, \boldsymbol{P}) &= \begin{vmatrix} -\frac{\partial T_k^1}{\partial u^1} & \cdots & -\frac{\partial T_k^1}{\partial u^N} \\ \vdots & \ddots & \vdots \\ -\frac{\partial T_k^N}{\partial u^1} & \cdots & -\frac{\partial T_k^N}{\partial u^N} \end{vmatrix} \\
&= \frac{1}{\beta_k^N}\frac{e^{Nz_k}\left(1+(N-1)e^{z_k}\right)^N}{\left(1+Ne^{z_k}\right)^{2N}} \begin{vmatrix} -1 & \frac{e^{z_k}}{1+(N-1)e^{z_k}} & \cdots & \frac{e^{z_k}}{1+(N-1)e^{z_k}} \\ \vdots & -1 & \ddots & \vdots \\ \frac{e^{z_k}}{1+(N-1)e^{z_k}} & \frac{e^{z_k}}{1+(N-1)e^{z_k}} & \cdots & -1 \end{vmatrix} \\
&= \frac{1}{\beta_k^N}\frac{e^{Nz_k}\left(-1\right)^N}{\left(1+Ne^{z_k}\right)^{N+1}} \neq 0.
\end{aligned}$$

(B.42)

We can now apply Lemma 2 and equations (B.37) to (B.41) into (B.8) to obtain

$$\boldsymbol{\beta z} = (\Phi - H(\boldsymbol{z}))\,\Omega(\boldsymbol{z}) - \boldsymbol{u}_0, \tag{B.43}$$

where $\boldsymbol{u}_0 = (u_b^0, u_s^0)$, $H(\boldsymbol{z})$ is a 2×2 matrix defined as

$$H(\boldsymbol{z}) := \begin{bmatrix} L_b d_b K_s + h_b - \phi_{bb} & -\phi_{sb}(d_s L_b + 1) \\ -\phi_{bs}(d_b L_s + 1) & L_s d_s K_b + h_s - \phi_{ss} \end{bmatrix}, \tag{B.44}$$

and, for any $k \in \{b, s\}$, L_k, d_k and h_k are functions depending on z_k as

$$\begin{aligned} L_k &= \frac{(N-1)\beta_k}{J_\phi}(1 + Ne^{z_k}), \\ d_k &= \beta_k(1 + Ne^{z_k}), \\ h_k &= \beta_k(1 + e^{z_k})(e^{-z_k} + N), \end{aligned} \tag{B.45}$$

and J_ϕ is a function of z_b and z_s as

$$\begin{aligned} K_k &= \phi_{kk} - \beta_k(1 + Ne^{z_k})(e^{-z_k} + N - 1), \quad \text{for } k \in \{b, s\}, \\ J_\phi &= K_b K_s - \phi_{sb}\phi_{bs}. \end{aligned} \tag{B.46}$$

Denoting $\boldsymbol{z}^*$ to be the solution to (B.43) and using (B.40) and (B.41), we conclude the proposition by noting that the symmetric equilibrium solution of (3.4) is given by $\boldsymbol{x}^* = \Omega(\boldsymbol{z}^*)$ and $\boldsymbol{p}^* = \Phi\Omega(\boldsymbol{z}^*) - \boldsymbol{\beta z}^* - \boldsymbol{u}^0$. □

Preliminary results for the proof of Proposition 10. We introduce notation and definitions and establish a useful lemma. Let $j \in \{1, \cdots, N\}$, k, l and $m \in \{b, s\}$, $u_{k,l} := \frac{\partial \phi_k}{\partial x_b} y_{b,l} + \frac{\partial \phi_k}{\partial x_s} y_{s,l}$ where $y_{k,l}$ is given by (B.21). Moreover,

$$\begin{aligned} U^j_{k,lm} :=& \frac{\partial}{\partial x^1_m}\left[\frac{\partial T^j_k}{\partial u_1}\right]\left(-\frac{\partial p^1_k}{\partial x^1_l} + \frac{\partial \phi_k}{\partial x_l}\right) + \sum_{i=2}^{N} \frac{\partial}{\partial x^1_m}\left[\frac{\partial T^j_k}{\partial u_i}\right] u_{k,l} + \\ & \sum_{i=2}^{N} \frac{\partial T^j_k}{\partial u_i}\left(\frac{\partial^2 \phi_k}{\partial x_m \partial x_b}\frac{\partial x^i_b}{\partial x^1_l} + \frac{\partial^2 \phi_k}{\partial x_m \partial x_s}\frac{\partial x^i_s}{\partial x^1_l}\right), \end{aligned} \tag{B.47}$$

where the derivatives of T_k^j are evaluated at $(u_k^0, u_k^1, \cdots, u_k^N)$, $u_k^1 := \phi_k(\boldsymbol{x}^1) - p_k^1$ and $u_k^i := \phi_k(\boldsymbol{x}^i) - p_k^*$ for $i \geq 2$. The following Lemma shows the second-order condition of (3.4) as a function of x_k.

Lemma 3 (SOC of CNE)**.** *The second-order condition of* (3.4) *is given by*

$$D^2_{(x_b^1, x_s^1)}\pi^1 = \left[\begin{array}{cc} 2\frac{\partial p_b^1}{\partial x_b^1} + x_b^1 \frac{\partial^2 p_b^1}{\partial \left(x_b^1\right)^2} + x_s^1 \frac{\partial^2 p_s^1}{\partial \left(x_b^1\right)^2} & \left(\frac{\partial p_b^1}{\partial x_s^1} + \frac{\partial p_s^1}{\partial x_b^1}\right) + x_b^1 \frac{\partial^2 p_b^1}{\partial x_s^1 \partial x_b^1} + x_s^1 \frac{\partial^2 p_s^1}{\partial x_s^1 \partial x_b^1} \\ \left(\frac{\partial p_s^1}{\partial x_b^1} + \frac{\partial p_b^1}{\partial x_s^1}\right) + x_b^1 \frac{\partial^2 p_b^1}{\partial x_b^1 \partial x_s^1} + x_s^1 \frac{\partial^2 p_s^1}{\partial x_b^1 \partial x_s^1} & 2\frac{\partial p_s^1}{\partial x_s^1} + x_b^1 \frac{\partial^2 p_b^1}{\partial (x_s^1)^2} + x_s^1 \frac{\partial^2 p_s^1}{\partial (x_s^1)^2} \end{array}\right], \quad \text{(B.48)}$$

where for k, m *and* $l \in \{b, s\}$, $\frac{\partial p_k^1}{\partial x_l^1}$ *is given by* (B.28)-(B.31),

$$\frac{\partial^2 p_k^1}{\partial x_m^1 \partial x_l^1} = \frac{\partial^2 \phi_k}{\partial x_m \partial x_l} + \frac{1}{J_k}\left(\left(S_k + (N-2)\, R_k\right) U_{k,lm}^1 + \left(x_{k,ml} - U_{k,lm}\right)(N-1)\, R_k\right), \quad \text{(B.49)}$$

and R_k, S_k *and* J_k *are given by* (B.7)*. Moreover,*

$$\left[\begin{array}{c} x_{b,ml} \\ x_{s,ml} \end{array}\right] = \frac{1}{J_\phi}\left[\begin{array}{cc} \left(\frac{\partial \phi_s}{\partial x_s} - \frac{1}{J_s} S_s\right) & -\frac{\partial \phi_b}{\partial x_s} \\ -\frac{\partial \phi_s}{\partial x_b} & \left(\frac{\partial \phi_b}{\partial x_b} - \frac{1}{J_b} S_b\right) \end{array}\right]\left[\begin{array}{c} \frac{1}{J_b}\left(R_b U_{b,lm}^1 - S_b U_{b,lm}\right) \\ \frac{1}{J_s}\left(R_s U_{s,lm}^1 - S_s U_{s,lm}\right) \end{array}\right], \quad \text{(B.50)}$$

where $U_{k,lm}^1$ *and* $U_{k,lm} := U_{k,lm}^j$ *for* $j \geq 2$ *are given by* (B.47).

The proof of this Lemma does not require assumptions I and II of Section 3.2. Thus, the SOC given by (B.48) is applicable to idiosyncratic preferences other than Gumbel distribution and to more general externality functions $\phi_k(\boldsymbol{x})$.

Proof of Lemma 3. Differentiating the left-hand side of (B.10) w.r.t. x_m^1, for $m \in \{b, s\}$, easily yields (B.48). To obtain (B.49) and (B.50), we differentiate (B.19) and (B.20) w.r.t. x_m^1. For $m, k, l \in \{b, s\}$ and $j \in \{2, 3, \cdots, N\}$, we obtain

$$\frac{\partial T_k^1}{\partial u_1}\left(-\frac{\partial^2 p_k^1}{\partial x_m^1 \partial x_l^1} + \frac{\partial^2 \phi_k}{\partial x_m^1 \partial x_l}\right) + \sum_{i=2}^N \frac{\partial T_k^1}{\partial u_i}\left(\frac{\partial \phi_k}{\partial x_b}\frac{\partial^2 x_b^i}{\partial x_m^1 \partial x_l^1} + \frac{\partial \phi_k}{\partial x_s}\frac{\partial^2 x_s^i}{\partial x_m^1 \partial x_l^1}\right) + U_{k,lm}^1 = 0, \quad \text{(B.51)}$$

$$\frac{\partial T_k^j}{\partial u_1}\left(-\frac{\partial^2 p_k^1}{\partial x_m^1 \partial x_l^1} + \frac{\partial^2 \phi_k}{\partial x_m^1 \partial x_l}\right) + \sum_{i=2}^N \frac{\partial T_k^j}{\partial u_i}\left(\frac{\partial \phi_k}{\partial x_b}\frac{\partial^2 x_b^i}{\partial x_m^1 \partial x_l^1} + \frac{\partial \phi_k}{\partial x_s}\frac{\partial^2 x_s^i}{\partial x_m^1 \partial x_l^1}\right) + U_{k,lm}^j = \frac{\partial^2 x_k^j}{\partial x_m^1 \partial x_l^1}, \quad \text{(B.52)}$$

where $U^1_{k,lm}$ and $U^j_{k,lm}$ are given by (B.47). Note that the derivatives of T^j_k in (B.51) and (B.52), are evaluated at $(u^0_k, u^1_k, \cdots, u^N_k)$, $u^1_k := \phi_k(\boldsymbol{x}^1) - p^1_k$, where $u^i_k := \phi_k(\boldsymbol{x}^i) - p^*_k$ for $i \geq 2$. The unknowns are $\frac{\partial^2 p^1_k}{\partial x^1_m \partial x^1_l}$ and $\frac{\partial^2 x^i_k}{\partial x^1_m \partial x^1_l}$ for $k, l, m \in \{b, s\}$ and $i \in \{2, \ldots, N\}$, adding up to $8 + 8(N-1) = 8N$. Similarly, the number of equations in the system of equations described in (B.51) and (B.52) is $8N$.

Next, we apply the property of symmetry at equilibrium, where we denote

$$\begin{aligned} &U^j_{k,lm} =: U_{k,lm} \;\text{ and} \\ &\frac{\partial^2 x^j_k}{\partial x^1_m \partial x^1_l} =: x_{k,ml}, \;\text{for} j \in \{2, \ldots, N\}. \end{aligned} \tag{B.53}$$

Incorporating (B.7), (B.21) and (B.53), we can reduce the system described by (B.51) and (B.52) into 16 equations with unknowns: $\frac{\partial^2 p^1_k}{\partial x^1_m \partial x^1_l}$ and $x_{k,ml}$ for $k, l, m \in \{b, s\}$. For $k, m, l \in \{b, s\}$, the 16 equations are

$$\begin{aligned} S_k q_k + (N-1)\, R_k t_k &= -U^1_{k,lm}, \\ R_k q_k + (S_k + (N-2)\, R_k)\, t_k &= x_{k,ml} - U_{k,lm}, \end{aligned} \tag{B.54}$$

where $q_k = \left(-\frac{\partial^2 p^1_k}{\partial x^1_m \partial x^1_l} + \frac{\partial^2 \phi_k}{\partial x^1_m \partial x_l}\right)$ and $t_k = \left(\frac{\partial \phi_k}{\partial x_b} x_{b,ml} + \frac{\partial \phi_k}{\partial x_s} x_{s,ml}\right)$. Solving the 2x2 system given by (B.54) easily yields (B.49) and (B.50). □

Lemma 3 is general enough to accommodate idiosyncratic preferences other than Gumbel distribution and general externality functions $\phi_k(\boldsymbol{x})$. Next, we use Assumptions I and II from Section 3.2, Proposition 9 and Lemma 3 to prove Proposition 10.

Proof of Proposition 10. The proof has two main steps: (i) Verifying sufficient conditions for (3.13) to have a unique solution; (ii) Establishing that a second order condition is satisfied.

Step (i): Note that (3.13) is equivalent to

$$\begin{aligned} &(2\phi_{bb} - \phi_{ss} d_b L_b + h_b)\, \omega\,(z_b) + (\phi_{sb} + \phi_{bs} + \phi_{sb} d_s L_b)\, \omega\,(z_s) - u^0_b = \beta_b z_b \;\text{ and} \\ &(\phi_{sb} + \phi_{bs} + \phi_{bs} d_b L_s)\, \omega\,(z_b) + (2\phi_{ss} - \phi_{bb} d_s L_s + h_s)\, \omega\,(z_s) - u^0_s = \beta_s z_s, \end{aligned} \tag{B.55}$$

where for any $k \in \{b, s\}$, L_k, d_k and h_k are functions depending on z_k given by (B.45). We want to find sufficient conditions for (B.55) to have a unique solution. First, we write some definitions. Set $\varphi_1 = (\phi_{bs}, \phi_{sb})$ and $\varphi_2 = (\phi_{bb}, \phi_{ss})$. For the given parameters in

$\psi = (\beta_b, \beta_s, u_b^0, u_s^0, N)$, we define $M : \mathbb{R}^6 \longrightarrow \mathbb{R}^2$ as

$$M\left(z_b, z_s, \varphi_1, \varphi_2; \psi\right) := \begin{bmatrix} M_b\left(z_b, z_s, \varphi_1, \varphi_2; \psi\right) \\ M_s\left(z_b, z_s, \varphi_1, \varphi_2; \psi\right) \end{bmatrix}, \text{ where for each } k, j \in \{b, s\},\ j \neq k,$$

$$M_k\left(z_b, z_s, \varphi_1, \varphi_2; \psi\right) := \left(2\phi_{kk} - \phi_{jj} d_k L_k + h_k\right) \omega\left(z_k\right) + \left(\phi_{sb} + \phi_{bs} + \phi_{jk} d_j L_k\right) \omega\left(z_j\right) - u_k^0 - \beta_k z_k. \tag{B.56}$$

We show that under (3.19), equation (B.55) has a unique solution for $\varphi_1 = 0$. From (B.56), if we let $\varphi_1 = 0$, we obtain

$$M_k\left(z_b, z_s, 0, \varphi_2; \psi\right) = \left(2\phi_{kk} - \phi_{jj} d_k L_k + h_k\right) \omega\left(z_k\right) - u_k^0 - \beta_k z_k. \tag{B.57}$$

Plugging (B.45) into (B.57) gives

$$M_k\left(z_b, z_s, 0, \varphi_2; \psi\right) = \frac{-\beta_k^2\left(1 + N e^{z_k}\right)^3 + \beta_k \phi_{kk} e^{z_k}\left(\left(2N - 1\right) e^{z_k} + 3\right)\left(1 + N e^{z_k}\right) - 2 e^{2 z_k} \phi_{kk}^2}{\left(1 + N e^{z_k}\right)\left(\beta_k\left(1 + \left(N - 1\right) e^{z_k}\right)\left(1 + N e^{z_k}\right) - e^{z_k} \phi_{kk}\right)} - \beta_k z_k - u_k^0. \tag{B.58}$$

It follows that for $\varphi_1 = 0$, M_k does not depend on z_j for $j \neq k$. Under (3.19), we claim that if $\varphi_1 = 0$, then the three statements below, (i-a)-(i-c), hold true:

(i-a) $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right)$ is continuous on z_k for all $z_k \in \mathbb{R}$.

(i-b) $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right)$ is strictly decreasing in the variable z_k for all $z_k \in \mathbb{R}$.

(i-c) The following limits hold true

$$\begin{aligned} &\lim_{z_k \to -\infty} M_k\left(z_b, z_s, 0, \varphi_2; \psi\right) = \infty \text{ and} \\ &\lim_{z_k \to \infty} M_k\left(z_b, z_s, 0, \varphi_2; \psi\right) = -\infty. \end{aligned} \tag{B.59}$$

Before proving the above claims, note that (i-a), (i-b) and (i-c) combined imply that there is a unique $(z_b^*, z_s^*) \in \mathbb{R}^2$ such that

$$M\left(z_b^*, z_s^*, 0, \varphi_2; \psi\right) = \begin{bmatrix} M_b\left(z_b^*, z_s^*, 0, \varphi_2; \psi\right) \\ M_s\left(z_b^*, z_s^*, 0, \varphi_2; \psi\right) \end{bmatrix} = 0.$$

Thus, under (3.19), equation (B.55) has a unique solution for $\varphi_1 = 0$. By (i-b),

$$\det\left(\frac{\partial\left(M_b, M_s\right)}{\partial\left(z_b, z_s\right)}\right)\Big|_{\left(z_b^*, z_s^*, 0, \varphi_2; \psi\right)} = \frac{\partial M_b\left(z_b^*, z_s^*, 0, \varphi_2; \psi\right)}{\partial z_b}\frac{\partial M_s\left(z_b^*, z_s^*, 0, \varphi_2; \psi\right)}{\partial z_s} > 0. \tag{B.60}$$

By (B.60) and the Implicit Function Theorem, there exists $\epsilon > 0$ and a unique continuous function

$$\left(z_b\left(\cdot\right), z_s\left(\cdot\right)\right) : B_\epsilon\left(0, 0\right) \longrightarrow \mathbb{R}^2$$

such that $\left(z_b\left(0, 0\right), z_s\left(0, 0\right)\right) = \left(z_b^*, z_s^*\right)$. Moreover, for all $\varphi_1 \in B_\epsilon\left(0, 0\right)$,

$$M\left(z_b\left(\varphi_1\right), z_s\left(\varphi_1\right), \varphi_1, \varphi_2; \psi\right) = 0.$$

In particular, under (3.19), there exists $\epsilon > 0$ such that for all $\varphi_1 \in B_\epsilon\left(0, 0\right)$, equation (B.55) has a unique solution. We now prove that under (3.19), if $\varphi_1 = 0$, then (i-a)-(i-c) hold true.

Proof of (i-a). We prove that $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right)$ is continuous on z_k for all $z_k \in \mathbb{R}$. The auxiliary function

$$g\left(z_k\right) := e^{-z_k}\left(1 + \left(N-1\right)e^{z_k}\right)\left(1 + Ne^{z_k}\right)$$

has a unique minimum at $z_k^0 = -\frac{1}{2}\ln(N\left(N-1\right))$, with $g\left(z_k^0\right) = 2\sqrt{N\left(N-1\right)} + 2N - 1$. It follows that if

$$\beta_k > \frac{\phi_{kk}}{\left(2\sqrt{N\left(N-1\right)} + 2N - 1\right)}, \tag{B.61}$$

then, the denominator of the fraction in (B.58) is never zero. We further prove (see the Mathematica file *Gumbel_N.nb*) that for all $\phi_{kk} > 0$,

$$f(N)\phi_{kk} \geq \frac{\phi_{kk}}{\left(2\sqrt{N\left(N-1\right)} + 2N - 1\right)} \tag{B.62}$$

(recall that $f(N)$ is stated in the proposition and given by (3.18)). Thus, if $\left(\phi_{kk}, \beta_k\right)$ satisfies (3.19), then (B.61) is satisfied and $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right)$ is continuous on z_k for all $z_k \in \mathbb{R}$.

Proof of (i-b). We prove that for each k, $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right)$ is strictly decreasing in the variable z_k for all $z_k \in \mathbb{R}$. In the supplementary file *Gumbel_N.nb*, we show that the partial derivative of $M_k(z_b, z_s, 0, \varphi_2; \psi)$ w.r.t. z_k can be written as

$$\frac{\partial M_k(z_b, z_s, 0, \varphi_2; \psi)}{\partial z_k} = -\frac{\sum_{m=0}^{6} a_m e^{m z_k}}{(1 + N e^{z_k})^2 (\beta_k (1 + (N-1) e^{z_k})(1 + N e^{z_k}) - e^{z_k} \phi_{kk})^2}, \tag{B.63}$$

where the coefficients $\{a_m\}_{m=0}^{6}$ are polynomials on the parameters $\{\phi_{kk}, \beta_k, N\}$ and are given by

$$\begin{aligned}
a_0 &= \beta_k^3, \\
a_1 &= \beta_k^2 (\beta_k (6N-1) - 4\phi_{kk}), \\
a_2 &= \beta_k (\beta_k^2 (15N^2 - 6N + 1) + 4\beta_k (1-4N) \phi_{kk} + 5\phi_{kk}^2), \\
a_3 &= 2N\beta_k^3 (10N^2 - 7N + 2) + \beta_k^2 \phi_{kk} (-24N^2 + 11N - 1) + \beta_k \phi_{kk}^2 (10N - 3) - 2\phi_{kk}^3, \\
a_4 &= \beta_k N (N\beta_k^2 (15N^2 - 16N + 6) + \beta_k \phi_{kk} (-16N^2 + 10N - 2) + (5N-2) \phi_{kk}^2), \\
a_5 &= \beta_k N^2 (N\beta_k^2 (6N^2 - 9N + 4) + \beta_k \phi_{kk} (-4N^2 + 3N - 1) + \phi_{kk}^2), \\
a_6 &= \beta_k^3 (N-1)^2 N^4.
\end{aligned} \tag{B.64}$$

Because (B.61) is satisfied, the denominator of (B.63) is always positive. We then show (see the Mathematica file *Gumbel_N.nb*) that under (3.19), i.e., if $N \geq 2$ and for each $k \in \{b, s\}$, (ϕ_{kk}, β_k) satisfies either

$$(\phi_{kk} \leq 0 \text{ and } \beta_k > 0) \text{ or } (\phi_{kk} > 0 \text{ and } \beta_k > f(N) \phi_{kk}),$$

then $a_m > 0$ for all $m = 0, \ldots, 6$, where $f(N)$ is given by (3.18). From (B.63), it follows that

$$\frac{\partial M_k(z_b, z_s, 0, \varphi_2; \psi)}{\partial z_k} < 0 \text{ for all } z_k. \tag{B.65}$$

It follows that $M_k(z_b, z_s, 0, \varphi_2; \psi)$ is strictly decreasing in the variable z_k for all $z_k \in \mathbb{R}$.

Proof of (i-c). This claim follows from applying L'hopital rule in (B.58).

Step (ii): We show that a second-order condition is satisfied. From (B.35), (B.36) and the substitution $z_k = \frac{u_k - u_k^0}{\beta_k}$, we obtain for $i, r, j \in \{1, \ldots, N\}$ and $k \in \{b, s\}$,

$$\frac{\partial^2 T_k^i\left(u_k^0, u_k, \ldots, u_k\right)}{\partial u^r \partial u^j} = \frac{e^{z_k}}{\left(1+Ne^{z_k}\right)^3 \beta_k^2} \cdot \begin{cases} \left(1+(N-1)\, e^{z_k}\right)\left(1+(N-2)\, e^{z_k}\right) & i=j=r \\ -\left(1+(N-2)\, e^{z_k}\right) e^{z_k} & i=j, i \neq r \\ -\left(1+(N-2)\, e^{z_k}\right) e^{z_k} & i \neq j, j \neq r, i=r \\ 2e^{2z_k} - \left(1+Ne^{z_k}\right) e^{z_k} & i \neq j = r \\ 2e^{2z_k} & i \neq j, j \neq r, i \neq r \end{cases} . \tag{B.66}$$

In the supplementary file *Gumbel_N.nb*, we show that the matrix $D^2_{(x_b^1, x_s^1)}\pi^1|_{\varphi_1=0}$, as given by (B.48), can be written as

$$D^2_{(x_b^1, x_s^1)}\pi^1\Big|_{\varphi_1=0} = \text{diag}\left(\frac{\partial^2 \pi^1}{\partial \left(x_b^1\right)^2}, \frac{\partial^2 \pi^1}{\partial \left(x_s^1\right)^2}\right)\Big|_{\varphi_1=0}, \tag{B.67}$$

where

$$\frac{\partial^2 \pi^1}{\partial \left(x_k^1\right)^2}\Big|_{\varphi_1=0} = \frac{\sum_{m=0}^{7} s_m e^{m z_k}}{e^{z_k}\left(\beta_k\left((N-1)\, e^{z_k}+1\right)\left(Ne^{z_k}+1\right) - e^{z_k}\phi_{kk}\right)^3}. \tag{B.68}$$

The coefficients $\{s_m\}_{m=0}^7$ are polynomials on the parameters $\{\phi_{kk}, \beta_k, N\}$ and are given by

$$s_0 = -\beta_k^4,$$

$$s_1 = \beta_k^3\left(5\phi_{kk} + \beta_k\left(1-7N\right)\right),$$

$$s_2 = -3\beta_k^2\left(\beta_k^2 N\left(7N-2\right) + \beta_k\phi_{kk}\left(2-9N\right) + 3\phi_{kk}^2\right),$$

$$s_3 = \beta_k\left(5\beta_k^3 N^2\left(3-7N\right) + 4\beta_k^2\left(N\left(15N-7\right)+1\right)\phi_{kk} + 3\beta_k\left(3-11N\right)\phi_{kk}^2 + 7\phi_{kk}^3\right),$$

$$s_4 = 5\beta_k^4 N^3\left(4-7N\right) + 2\beta_k^3 N\left(N\left(35N-26\right)+7\right)\phi_{kk} + \beta_k^2\left(\left(26-45N\right)N-5\right)\phi_{kk}^2$$
$$+\beta_k\left(13N-4\right)\phi_{kk}^3 - 2\phi_{kk}^4,$$

$$s_5 = \beta_k\left(3\beta_k^3\left(5-7N\right)N^4 + 3\beta_k^2\left(N\left(15N-16\right)+6\right)N^2\phi_{kk} + \beta_k\left(\left(25-27N\right)N-10\right)N\phi_{kk}^2\right)$$

$$+\beta_k\left(\left(6N^2-4N+1\right)\phi_{kk}^3\right),$$

$$s_6=\beta_k N\left(\beta_k^3\left(6-7N\right)N^4+\beta_k^2\left(N(15N-22)+10\right)N^2\phi_{kk}+\beta_k\left(-6N^2+8N-5\right)N\phi_{kk}^2+\phi_{kk}^3\right),$$

$$s_7=-\beta_k^3\left(N-1\right)N^4\left(\beta_k N^2+2\left(-N+1\right)\phi_{kk}\right). \quad \text{(B.69)}$$

Because (B.61) is satisfied, the denominator of (B.68) is always positive. We then show (see the Mathematica file *Gumbel_N.nb*) that under (3.19), i.e., if $N\geq 2$ and for each $k\in\{b,s\}$, (ϕ_{kk},β_k) satisfies either

$$\left(\phi_{kk}\leq 0 \text{ and } \beta_k>0\right) \text{ or } \left(\phi_{kk}>0 \text{ and } \beta_k>f\left(N\right)\phi_{kk}\right),$$

then $s_m<0$ for all $m=0,\ldots,6$, where $f(N)$ is given by (3.18). It follows that $D^2_{(x_b^1,x_s^1)}\pi^1|_{\varphi_1=0}$ is negative definite. By continuity there exists $\tilde{\epsilon}>0$ such that for all $\varphi_1\in B_{\tilde{\epsilon}}\left(0,0\right)$,

$$D^2_{(x_b^1,x_s^1)}\pi^1\left(z_b\left(\varphi_1\right),z_s\left(\varphi_1\right),\varphi_1,\varphi_2;\psi\right)$$

is negative definite. Therefore, a second condition for (B.43) is satisfied.

□

Preliminary result for the proof of Proposition 11. Using the same notation we introduced in (B.7), the following lemma provides the FOC of (3.5) as a function of x_k.

Lemma 4 (FOC of CE)**.** *The symmetric collusive equilibrium* $\mathbf{p}^C$ *and* $\boldsymbol{x}^C$ *are solutions of* (3.3) *and of the following two equations*

$$p_k+x_k\left(\frac{\partial\phi_k}{\partial x_k}-\frac{1}{S_k+(N-1)R_k}\right)+x_l\frac{\partial\phi_l}{\partial x_k}=0,\quad \text{for } k,l\in\{b,s\},k\neq l\quad . \quad \text{(B.70)}$$

The proof of this Lemma does not require assumptions I and II in Section 3.2. Thus, the FOC given by (B.70) is applicable to idiosyncratic preferences other than Gumbel distribution and to more general externality functions $\phi_k(\boldsymbol{x})$.

Proof of Lemma 4. The FOC of (3.5) w.r.t. x_k is

$$\frac{\partial\Pi_{\text{tot}}}{\partial x_k}=N\left(p_k+x_k\frac{\partial p_k}{\partial x_k}+x_l\frac{\partial p_l}{\partial x_k}\right)=0,\quad \text{for } k,l\in\{b,s\},k\neq l. \quad \text{(B.71)}$$

To solve (B.71), we need the expressions of

$$\frac{\partial p_k}{\partial x_l}, \quad \text{for each } k, l \in \{b, s\}. \tag{B.72}$$

We determine those four partial derivatives $\frac{\partial \boldsymbol{p}^1}{\partial \boldsymbol{x}^1}$ in (B.72) using the definition of T_k in (3.8). By (3.9), for $k \in \{b, s\}$, the vectors of market shares and prices, $(x_k, \dots, x_k)$ and $(p_k, \dots, p_k)$ satisfy all the following

$$T_k^i(u_k^0, u_k, u_k, \dots, u_k) = x_k, \quad \text{for } i \in \{1, 2, \cdots, N\}. \tag{B.73}$$

Using the relationship $u_k = \phi_k(\boldsymbol{x}) - p_k$, it follows that

$$\frac{\partial u_k}{\partial x_l} = \frac{\partial \phi_k}{\partial x_l} - \frac{\partial p_k}{\partial x_l}.$$

Taking derivative w.r.t. x_b and x_s in (B.73) gives us

$$\left(\sum_{j=1}^{N} \frac{\partial T_k^i}{\partial u^j}\right)\left(\frac{\partial \phi_k}{\partial x_l} - \frac{\partial p_k}{\partial x_l}\right) = \delta_{kl}, \text{ for } i \in \{1, \cdots, N\}, \ k, l \in \{b, s\}, \tag{B.74}$$

where $\delta_{kl} = 1$ when $k = l$ and $\delta_{kl} = 0$ when $k \neq l$. From (B.7) and (B.74),

$$\frac{\partial p_k}{\partial x_l} = \frac{\partial \phi_k}{\partial x_l} - \frac{\delta_{kl}}{S_k + (N-1)R_k}. \tag{B.75}$$

Plugging (B.74) into (B.71) gives us

$$p_k + x_k\left(\frac{\partial \phi_k}{\partial x_k} - \frac{1}{S_k + (N-1)R_k}\right) + x_l \frac{\partial \phi_l}{\partial x_k} = 0, \quad \text{for } k, l \in \{b, s\}, k \neq l. \tag{B.76}$$

□

Lemma 4 is general enough to accommodate idiosyncratic preferences other than Gumbel distribution and general externality functions $\phi_k(\boldsymbol{x})$. Next, we use Assumptions I and II from Section 2 and Lemma 4 to prove Proposition 11.

Proof of Proposition 11. We want to rewrite the FOC given by (B.70) using Assumptions I and II from Section 3.2. We plugging (B.38), (B.39), (B.40) and (B.41) from the proof of

Proposition 9 into (B.70) to obtain

$$\boldsymbol{\beta z} = \left(\Phi - H^C(\boldsymbol{z})\right)\Omega(\boldsymbol{z}) - \boldsymbol{u}_0, \tag{B.77}$$

where $H^C(\boldsymbol{z})$ is a 2×2 matrix defined as

$$H(\boldsymbol{z}) := \begin{bmatrix} \frac{\beta_b(1+Ne^{z_b})^2}{e^{z_b}} - \phi_{bb} & -\phi_{sb} \\ -\phi_{bs} & \frac{\beta_s(1+Ne^{z_s})^2}{e^{z_s}} - \phi_{ss} \end{bmatrix}. \tag{B.78}$$

Denoting $\boldsymbol{z}^C$ to be the solution to (B.77) and using (B.41), we conclude the proposition by noting that the symmetric equilibrium solution of (3.5) is given by $\boldsymbol{x}^C = \Omega(\boldsymbol{z}^C)$ and $\boldsymbol{p}^C = \Phi\Omega(\boldsymbol{z}^C) - \boldsymbol{\beta z}^C - \boldsymbol{u}_0$. □

Preliminary results for the proof of Proposition 12. The following Lemma shows the second-order condition of (3.5) as a function of x_k.

Lemma 5 (SOC of CNE)**.** *The second-order condition of* (3.5) *is given by*

$$\frac{\partial^2 \Pi_{tot}}{\partial x_m \partial x_k} = N\left(\frac{\partial p_k}{\partial x_m} + \delta_{km}\frac{\partial p_k}{\partial x_k} + \delta_{ml}\frac{\partial p_l}{\partial x_k} + x_k\frac{\partial^2 p_k}{\partial x_m \partial x_k} + x_l\frac{\partial^2 p_l}{\partial x_m \partial x_k}\right), \tag{B.79}$$

for $k, l, m \in \{b, s\}, k \neq l$, *where* $\frac{\partial p_k}{\partial x_l}$ *is given by* (B.75),

$$\frac{\partial^2 p_k}{\partial x_m \partial x_l} = \frac{\partial^2 \phi_k}{\partial x_m \partial x_l} + \frac{\delta_{km}\delta_{kl}}{\left(S_k + (N-1)R_k\right)^3}\left(\sum_{j=1}^{N}\sum_{r=1}^{N}\frac{\partial^2 T_k^i\left(u_k^0, u_k, \cdots, u_k\right)}{\partial u^r \partial u^j}\right), \tag{B.80}$$

for $k, l, m \in \{b, s\}$, *and* S_k, R_k *are given by* (B.7).

The proof of this Lemma does not require assumptions I and II of Section 3.2. Thus, the SOC given by (B.79) is applicable to idiosyncratic preferences other than Gumbel distribution and to more general externality functions $\phi_k(\boldsymbol{x})$.

Proof of Lemma 5. Differentiating the left-hand side of (B.71) w.r.t. x_m, for $m \in \{b, s\}$, easily yields (B.79). To obtain (B.80), we differentiate (B.74) w.r.t. x_m. For $m, k, l \in$

$\{b, s\}$, we obtain

$$\sum_{j=1}^{N}\sum_{r=1}^{N}\frac{\partial^2 T_k^i}{\partial u^r \partial u^j}\left(\frac{\partial \phi_k}{\partial x_m}-\frac{\partial p_k}{\partial x_m}\right)\left(\frac{\partial \phi_k}{\partial x_l}-\frac{\partial p_k}{\partial x_l}\right)+\left(\sum_{j=1}^{N}\frac{\partial T_k^i}{\partial u^j}\right)\left(\frac{\partial^2 \phi_k}{\partial x_m \partial x_l}-\frac{\partial^2 p_k}{\partial x_m \partial x_l}\right) = 0, \tag{B.81}$$

where the derivatives of T_k^i are evaluated at $(u_k^0, u_k, ..., u_k)$ and $u_k = \phi_k(\boldsymbol{x}) - p_k$. Plugging (B.7) and (B.75) into (B.81), yields (B.80). □

Lemma 5 is general enough to accommodate idiosyncratic preferences other than Gumbel distribution and general externality functions $\phi_k(\boldsymbol{x})$. Next, we use Assumptions I and II from Section 3.2, Proposition 11 and Lemma 5 to prove Proposition 12

Proof of Proposition 12. The proof has two main steps: (i) Verifying sufficient conditions for (3.20) to have a unique solution; (ii) Establishing that a second order condition is satisfied.

Step (i): Note that (3.20) is equivalent to

$$\begin{aligned} &\frac{2\phi_{bb}}{\beta_b(N+e^{-z_b})}+\frac{\phi_{bs}+\phi_{sb}}{\beta_b(N+e^{-z_s})}-\frac{u_b^0}{\beta_b}-(1+Ne^{z_b})=z_b \quad \text{and} \\ &\frac{\phi_{bs}+\phi_{sb}}{\beta_s(N+e^{-z_b})}+\frac{2\phi_{ss}}{\beta_s(N+e^{-z_s})}-\frac{u_s^0}{\beta_s}-(1+Ne^{z_s})=z_s. \end{aligned} \tag{B.82}$$

For each $k \in \{b, s\}$, $k \neq l$, we denote

$$F_k(z_b, z_s) := \frac{2\phi_{kk}}{\beta_k(N+e^{-z_k})}+\frac{\phi_{bs}+\phi_{sb}}{\beta_k(N+e^{-z_l})}-\frac{u_k^0}{\beta_k}-(1+Ne^{z_k}). \tag{B.83}$$

We want to find sufficient conditions for (B.82) to have a unique solution. By bounding each term of $F_k(z_b, z_s)$ independently, we can identify an upper bound for $F_k(z_b, z_s)$ in $\mathbb{R}^2$, indeed,

$$F_k(z_b, z_s) \leq \frac{|2\phi_{kk}|}{\beta_k N}+\frac{|\phi_{bs}+\phi_{sb}|}{\beta_k N}-\frac{u_k^0}{\beta_k}-1 := v_k. \tag{B.84}$$

Similarly, if $z_k \leq v_k$, a lower bound for $F_k(z_b, z_s)$ is

$$F_k(z_b, z_s) \geq -\frac{|2\phi_{kk}|}{\beta_k N} - \frac{|\phi_{bs} + \phi_{sb}|}{\beta_k N} - \frac{u_k^0}{\beta_k} - (1 + Ne^{v_k}) := w_k. \tag{B.85}$$

We denote a vector-valued function $\boldsymbol{F}(z_b, z_s) := (F_b(z_b, z_s), F_s(z_b, z_s))^T$. By combining (B.84) and (B.85), we conclude that $\boldsymbol{F}(z_b, z_s)$ maps the area $[w_b, v_b] \times [w_s, v_s]$ into $[w_b, v_b] \times [w_s, v_s]$. It is clear that $\boldsymbol{F}(z_b, z_s)$ is a continuous function. By using Brouwer's Fixed-Point Theorem, there is a fixed point for the function $\boldsymbol{F}(z_b, z_s)$ in the area $[w_b, v_b] \times [w_s, v_s]$, and this concludes that there is a solution for (3.20) in the area $[w_b, v_b] \times [w_s, v_s]$.

We now prove the uniqueness of the solution. We first consider $\phi_{bs} = \phi_{sb} = 0$, then (3.20) becomes two decoupled equations. In particular, for each $k \in \{b, s\}$, z_k^{C} is the solution to

$$M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi) := \frac{2\phi_{kk}}{N + e^{-z_k}} - \beta_k (1 + Ne^{z_k}) - u_k^0 - \beta_k z_k = 0, \tag{B.86}$$

where we are setting $\varphi_1 = (\phi_{bs}, \phi_{sb}) = 0$, $\varphi_2 = (\phi_{bb}, \phi_{ss})$ and $\psi = (\beta_b, \beta_s, u_b^0, u_s^0, N)$. From (B.86), the function $M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi)$ is continuous for all $z_k \in \mathbb{R}$. Moreover,

$$\lim_{z_k \to -\infty} M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi) = \infty \quad \text{and} \quad \lim_{z_k \to \infty} M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi) = -\infty. \tag{B.87}$$

The partial derivative of $M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi)$ w.r.t. z_k can be written as

$$\frac{\partial M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi)}{\partial z_k} = \frac{2e^{z_k}\phi_{kk} - \beta_k (Ne^{z_k} + 1)^3}{(Ne^{z_k} + 1)^2}. \tag{B.88}$$

Given that $\beta_k > 0$, if $\phi_{kk} \leq 0$, then $M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi)$ is strictly decreasing w.r.t. z_k for all $z_k \in \mathbb{R}$. Now, suppose that $\phi_{kk} > 0$. The function $z_k \mapsto 2e^{z_k} / (Ne^{z_k} + 1)^3$ has a unique maximum over $\mathbb{R}$ at $z_k^0 = \log \frac{1}{2N}$ and such maximum is given by $\frac{8}{27N}$. Therefore, the numerator of (B.88) is strictly negative whenever $\beta_k > \frac{8\phi_{kk}}{27N}$. Thus, if either $\phi_{kk} \leq 0$ or ($\phi_{kk} > 0$ and $\beta_k > \frac{8\phi_{kk}}{27N}$), then

$$\frac{\partial M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi)}{\partial z_k} < 0 \quad \text{for all } z_k. \tag{B.89}$$

It follows that there is a unique solution for (B.86). Now, notice that

$$\det\left(\frac{\partial(M_b^C, M_s^C)}{\partial(z_b, z_s)}\right)\Big|_{(z_b^{\mathrm{C}}, z_s^{\mathrm{C}}, 0, \varphi_2;\psi)} = \frac{\partial M_b^{\mathrm{C}}\left(z_b^{\mathrm{C}}, z_s^{\mathrm{C}}, 0, \varphi_2; \psi\right)}{\partial z_b}\frac{\partial M_s^{\mathrm{C}}\left(z_b^{\mathrm{C}}, z_s^{\mathrm{C}}, 0, \varphi_2; \psi\right)}{\partial z_s} > 0. \tag{B.90}$$

By the Implicit Function Theorem, there exists $\epsilon > 0$ and a unique continuous function

$$(z_b^{\mathrm{C}}(\varphi_1), z_s^{\mathrm{C}}(\varphi_1)) : B_\epsilon(0,0) \longrightarrow \mathbb{R}^2$$

such that $(z_b^{\mathrm{C}}(0,0), z_s^{\mathrm{C}}(0,0)) = (z_b^{\mathrm{C}}, z_s^{\mathrm{C}})$. Moreover, for all $\varphi_1 \in B_\epsilon(0,0)$,

$$M^{\mathrm{C}}(z_b^C(\varphi_1), z_s^C(\varphi_1), \varphi_1, \varphi_2; \psi) := \begin{bmatrix} M_b^{\mathrm{C}}(z_b^C(\varphi_1), z_s^C(\varphi_1), \varphi_1, \varphi_2; \psi) \\ M_s^{\mathrm{C}}(z_b^C(\varphi_1), z_s^C(\varphi_1), \varphi_1, \varphi_2; \psi) \end{bmatrix} = 0.$$

Step (ii): We show that a second-order condition is satisfied. Combining (B.66), (B.79) and (B.80), we show (supplementary file *Gumbel_N.nb*) that

$$D^2_{(x_b^1, x_s^1)}\Pi_{tot} =$$
$$\begin{bmatrix} -Ne^{-z_b}\left(\beta_b\left(Ne^{z_b}+1\right)^3 - 2e^{z_b}\phi_{bb}\right) & N(\phi_{bs}+\phi_{sb}) \\ N(\phi_{bs}+\phi_{sb}) & -Ne^{-z_s}\left(\beta_s\left(Ne^{z_s}+1\right)^3 - 2e^{z_s}\phi_{ss}\right) \end{bmatrix}. \tag{B.91}$$

Following the same argument from the paragraph above (B.89), if $\phi_{bs} = \phi_{sb} = 0$, $N \geq 2$ and for each $k \in \{b, s\}$, (ϕ_{kk}, β_k) satisfies either

$$(\phi_{kk} \leq 0 \text{ and } \beta_k > 0) \text{ or } \left(\phi_{kk} > 0 \text{ and } \beta_k > \frac{8\phi_{kk}}{27N}\right),$$

then $D^2_{(x_b^1, x_s^1)}\Pi_{tot}|_{\varphi_1=0}$ is negative definite. By continuity there exists $\tilde{\epsilon} > 0$ such that for all $\varphi_1 \in B_{\tilde{\epsilon}}(0,0)$,

$$D^2_{(x_b^1, x_s^1)}\Pi_{tot}\left(z_b\left(\varphi_1\right), z_s\left(\varphi_1\right), \varphi_1, \varphi_2; \psi\right)$$

is negative definite. Therefore, a second condition for (B.77) is satisfied.

□

Proof of Proposition 13. Suppose that $N \geq 2$ and for each $k \in \{b, s\}$, (ϕ_{kk}, β_k) satisfies (3.19). Assume that $\varphi_1 = (\phi_{bs}, \phi_{sb}) = 0$. From the proof of Proposition 10, M_k does not depend on z_j for $j \neq k$ (see (B.58)). Moreover,

$$\lim_{z_k \to -\infty} M_k\left(z_b, z_s, 0, \varphi_2; \psi\right) = \infty, \tag{B.92}$$

and $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right)$ is strictly decreasing in z_k for all $z_k \in \mathbb{R}$. From (B.58), we compute

$$\begin{aligned} M_k\left(z_b, z_s, 0, \varphi_2; \psi\right)\Big|_{z_k=0} &= -\frac{\beta_k^2\left(N+1\right)^3 - 2\phi_{kk}\beta_k\left(N+1\right)^2 + 2\phi_{kk}^2}{\left(N+1\right)\left(\beta_k N\left(N+1\right) - \phi_{kk}\right)} - u_k^0 \\ &= -\frac{A_k}{\left(N+1\right)\left(\beta_k N\left(N+1\right) - \phi_{kk}\right)}, \end{aligned} \tag{B.93}$$

where A_k is a polynomial of order 2 in β_k given by

$$A_k := \beta_k^2\left(N+1\right)^3 + \beta_k\left(N+1\right)^2\left(Nu_k^0 - 2\phi_{kk}\right) - \phi_{kk}\left(\left(N+1\right)u_k^0 - 2\phi_{kk}\right). \tag{B.94}$$

The largest root of A_k is

$$\gamma(N, \phi_{kk}, u_k^0) := \frac{\left(2\phi_{kk} - Nu_k^0\right) + \sqrt{\left(2\phi_{kk} - Nu_k^0\right)^2 + 4\phi_{kk}\left(u_k^0 - \frac{2\phi_{kk}}{N+1}\right)}}{2\left(N+1\right)}. \tag{B.95}$$

We also denote the smallest root of A_k by γ_-. Note that by (3.19), the denominator of (B.93) is always positive.

Step (i): If $\beta_k > \gamma(N, \phi_{kk}, u_k^0)$, by (B.93), (B.94) and (B.95), then (B.93) is strictly negative. From (B.92), the fact that $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right)$ is strictly decreasing in z_k for all $z_k \in \mathbb{R}$, and that z_k^* is the unique solution of $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right) = 0$, then $z_k^* < 0$.

Step (ii): If $\gamma_- < \beta_k < \gamma(N, \phi_{kk}, u_k^0)$, then (B.93) is strictly positive. From (B.92), the fact that $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right)$ is strictly decreasing in z_k for all $z_k \in \mathbb{R}$, and that z_k^* is the unique solution of $M_k\left(z_b, z_s, 0, \varphi_2; \psi\right) = 0$, then $z_k^* > 0$.

Note that from the definition of γ_- as the smallest root of the second degree polynomial A_k, if $\phi_{kk} \leq 0$, then $\gamma_- \leq 0$. If $\phi_{kk} > 0$, then $f(N)\phi_{kk} > \gamma_-$ (see the supplementary file *Gumbel_N.nb*). It follows that (3.19) combined with $\beta_k < \gamma(N, \phi_{kk}, u_k^0)$ imply that $z_k^* > 0$.

To end the proof of this proposition, note that from the Proof of Proposition 10, there is $\epsilon > 0$ and a unique continuous function

$$\left(z_b^*\left(\cdot\right), z_s^*\left(\cdot\right)\right) : B_\epsilon\left(0,0\right) \longrightarrow \mathbb{R}^2$$

such that for all $\varphi_1 \in B_\epsilon(0,0)$, $M\left(z_b^*(\varphi_1), z_s^*(\varphi_1), \varphi_1, \varphi_2; \psi\right) = 0$. By continuity, for $\hat{\varepsilon} = |z_k^*(0,0)|/2$, there exists $\delta > 0$ such that for all $\varphi_1 \in B_{\min(\delta,\varepsilon)}(0,0)$,

$$z_k^*(0,0) - \frac{|z_k^*(0,0)|}{2} < z_k^*(\varphi_1) < z_k^*(0,0) + \frac{|z_k^*(0,0)|}{2}.$$

Thus, when $z_k^*(0,0) < 0$, by step (i) we obtain $z_k^*(\varphi_1) < \frac{z_k^*(0,0)}{2} < 0$. When $z_k^* > 0$, by (ii) we obtain $z_k^*(\varphi_1) > \frac{z_k^*(0,0)}{2} > 0$.

□

Proof of Corollary 1. We show that case (ii) in Proposition 13 is not feasible for large values of u_k^0. We first assume $\phi_{kk} \leq 0$ and $\beta_k > 0$. From (B.95),

$$\operatorname{sign}\left(\gamma\left(N, \phi_{kk}, u_k^0\right)\right) = \operatorname{sign}\left(\frac{2\phi_{kk}}{N+1} - u_k^0\right). \tag{B.96}$$

Set $u_{k,1}^0 := 2\phi_{kk}/(N+1)$. It follows that if $\phi_{kk} \leq 0$ and $u_k^0 \geq u_{k,1}^0$, then $\gamma\left(N, \phi_{kk}, u_k^0\right) \leq 0$. Thus, (ii) in Proposition 13 is not feasible as $\beta_k > 0$.

Next, we assume that $\phi_{kk} > 0$ and $\beta_k > f(N)\,\phi_{kk}$. We verify (see *Gumbel_N.nb*) that for all $\phi_{kk} > 0$ and $N \geq 2$,

$$\begin{cases} \gamma\left(N, \phi_{kk}, u_k^0\right) \leq f(N)\,\phi_{kk} & \text{for any } u_k^0 \geq u_{k,2}^0, \\ f(N)\,\phi_{kk} < \gamma\left(N, \phi_{kk}, u_k^0\right) & \text{for any } u_k^0 < u_{k,2}^0, \end{cases} \tag{B.97}$$

where

$$u_{k,2}^0 := -\frac{2\left(N^4 - 2N^3 - 2N^2 + 2N + 2\right)\phi_{kk}}{N^3\left(2N^3 + N^2 - 3N - 2\right)}. \tag{B.98}$$

In particular, if $u_k^0 \geq u_{k,2}^0$, then case (ii) in Proposition 13 is not feasible as $\beta_k > f(N)\,\phi_{kk}$.

We finish the proof with a piecewise definition for $\tilde{u}_k^0$,

$$\tilde{u}_k^0 := \begin{cases} u_{k,1}^0 & \text{if } \phi_{kk} \leq 0, \\ u_{k,2}^0 & \text{if } \phi_{kk} > 0, \end{cases} \tag{B.99}$$

where $u_{k,1}^0 = 2\phi_{kk}/(N+1)$ and $u_{k,2}^0$ is given by (B.98).

□

Proof of Corollary 2. For each $k \in \{b, s\}$, let $u_k^0 \in \mathbb{R}$, $\Phi \in \mathbb{R}^{2x2}$ and $\beta_k > 0$. From (B.55), as $N \to \infty$, the FOC of (3.4) becomes

$$-u_k^0 - \beta_k (z_k + 1) = 0, \text{ for each } k \in \{b, s\}.$$

Thus, $\lim_{N\to\infty} z_k^* = -\left(\frac{u_k^0}{\beta_k} + 1\right)$. Moreover, $\lim_{N\to\infty} N x_k^* = 1$ and $\lim_{N\to\infty} p_k^* = \beta_k$. The solution of the equation $z_k^* = 0$ when $N \to \infty$ is $\beta_k = -u_k^0$ if $u_k^0 < 0$. If $u_k^0 \geq 0$, then $\lim_{N\to\infty} z_k^* < 0$.

□

Proof of Proposition 14. Recall from (B.41) in the Proof of Proposition 9,

$$\boldsymbol{p}^* = \Phi\Omega(\boldsymbol{z}^*) - \boldsymbol{\beta}\boldsymbol{z}^* - \boldsymbol{u}_0,$$

where $\Omega(\boldsymbol{z}^*) = (\omega(z_b^*), \omega(z_s^*))^T$, $\omega(\cdot)$ and $\boldsymbol{z}^*$ are given by (B.40) and (B.55), respectively. We want to compute the following quantity when $\varphi_1 = 0$,

$$\left.\frac{\partial p_k^*}{\partial u_k^0}\right|_{\varphi_1=0} = \left[\frac{\phi_{kk} e^{z_k^*}}{\left(N e^{z_k^*} + 1\right)^2} - \beta_k\right] \frac{\partial z_k^*}{\partial u_k^0} - 1, \quad k \in \{b, s\}. \tag{B.100}$$

By (B.58), z_k^* is uniquely characterized by $M_k(z_b^*, z_s^*, 0, \varphi_2; \psi) + u_k^0 - u_k^0 = 0$. It follows that

$$\frac{\partial z_k^*}{\partial u_k^0} = \left[\frac{\partial\left(M_k(z_b^*, z_s^*, 0, \varphi_2; \psi) + u_k^0\right)}{\partial z_k}\right]^{-1} = \left[\frac{\partial M_k(z_b^*, z_s^*, 0, \varphi_2; \psi)}{\partial z_k}\right]^{-1}, \tag{B.101}$$

where $\frac{\partial M_k\left(z_b^*, z_s^*, 0, \varphi_2; \psi\right)}{\partial z_k}$ is given by (B.63). After plugging (B.101) into (B.100), we show (see *Gumbel_N.nb*) that

$$\left.\frac{\partial p_k^*}{\partial u_k^0}\right|_{\varphi_1=0} = -\frac{n_{p,u}(z_k^*, \beta_k, \phi_{kk}, N)}{d_{p,u}(z_k^*, \beta_k, \phi_{kk}, N)}, \tag{B.102}$$

where $n_{p,u}(z_k^*, \beta_k, \phi_{kk}, N)$ and $d_{p,u}(z_k^*, \beta_k, \phi_{kk}, N)$ can be written as polynomials in $e^{z_k^*}$ in the following way,

$$n_{p,u}\left(z_k^*,\beta_k,\phi_{kk},N\right) := \sum_{m=1}^{5} n_{p,u,m} e^{m z_k^*}, \quad \text{and}$$
$$d_{p,u}\left(z_k^*,\beta_k,\phi_{kk},N\right) := \sum_{m=0}^{6} d_{p,u,m} e^{m z_k^*}. \tag{B.103}$$

Moreover, the coefficients $n_{p,u,m}$ are as follows:

$$\begin{aligned}
n_{p,u,1} &= \beta_k^2\left(\beta_k - \phi_{kk}\right), \\
n_{p,u,2} &= 2\beta_k\left(2\beta_k^2 N - 2\beta_k N\phi_{kk} + \phi_{kk}^2\right), \\
n_{p,u,3} &= 6\beta_k^3 N^2 - N\beta_k^2\left(6N+1\right)\phi_{kk} + \phi_{kk}^2\beta_k\left(4N-1\right) - \phi_{kk}^3, \\
n_{p,u,4} &= 2\beta_k N^2\left(\beta_k - \phi_{kk}\right)\left(2\beta_k N - \phi_{kk}\right), \text{ and} \\
n_{p,u,5} &= \beta_k N^2\left(\beta_k^2 N^2 - \phi_{kk} N\beta_k\left(N+1\right) + \phi_{kk}^2\right).
\end{aligned} \tag{B.104}$$

The coefficients $d_{p,u,m}$ are given by

$$d_{p,u,0} = \beta_k^3,$$

$$d_{p,u,1} = \beta_k^2\left(\beta_k\left(6N-1\right) - 4\phi_{kk}\right),$$

$$d_{p,u,2} = \beta_k\left(\beta_k^2\left(15N^2-6N+1\right) + 4\beta_k\left(1-4N\right)\phi_{kk} + 5\phi_{kk}^2\right),$$

$$d_{p,u,3} = N\beta_k^3\left(20N^2-14N+4\right) + \beta_k^2\left(-24N^2+11N-1\right)\phi_{kk} + \beta_k\left(10N-3\right)\phi_{kk}^2 - 2\phi_{kk}^3,$$

$$d_{p,u,4} = \beta_k N\left(N\beta_k^2\left(15N^2-16N+6\right) + \beta_k\left(-16N^2+10N-2\right)\phi_{kk} + \left(5N-2\right)\phi_{kk}^2\right),$$

$$d_{p,u,5} = \beta_k N^2\left(N\beta_k^2\left(6N^2-9N+4\right) + \beta_k\left(-4N^2+3N-1\right)\phi_{kk} + \phi_{kk}^2\right), \text{ and}$$

$$d_{p,u,6} = \beta_k^3\left(N-1\right)^2 N^4. \tag{B.105}$$

Because the expressions determining $n_{p,u}\left(z_k^*,\beta_k,\phi_{kk},N\right)$ and $d_{p,u}\left(z_k^*,\beta_k,\phi_{kk},N\right)$ are complex, we focus on finding sufficient conditions for these expressions to have a specific sign for all z_k^*.

Case (i): $\left.\frac{\partial p_k^*}{\partial u_k^0}\right|_{\varphi_1=0} < 0$. We verify in the supplementary file *Gumbel_N.nb* that $n_{p,u}$ and $d_{p,u}$ (see (B.103)) are positive, if either of the two conditions below, (i-a) or (i-b), hold:

(i-a) $\phi_{kk} \leq 0$, $N \geq 2$ and $\beta_k > 0$.

(i-b) $\phi_{kk} > 0$, $N \geq 2$ and $\beta_k > g_{p,u}(N, \phi_{kk})$, where $g_p(N, \phi_{kk})$ is the largest real root of the third degree polynomial $n_{p,u,5}$ (viewed as a polynomial in β_k).

Using the quadratic formula, we verify that $n_{p,u,5}$ (see (B.104)) has three real roots and that $g_{p,u}(N, \phi_{kk})$ is linear in ϕ_{kk} and can thus be expressed as $g_{p,u}(N, \phi_{kk}) = g_{p,u}(N)\,\phi_{kk}$ where

$$g_{p,u}(N) := \frac{\left(N + \sqrt{(N-1)(N+3)} + 1\right)}{2N}.$$

Case (ii): $\left.\frac{\partial p_k^*}{\partial N}\right|_{\varphi_1=0} > 0$. We verify in the supplementary file *Gumbel_N.nb* that $n_{p,u}$ and $d_{p,u}$ (see (B.103)) are negative and positive, respectively, if the condition below, (ii), holds:

(ii) $\phi_{kk} > 0$, $N \geq 3$ and $f(N)\phi_{kk} < \beta_k < f_{p,u}(N, \phi_{kk})$, where $f_{p,u}(N, \phi_{kk})$ is the largest real root of the third degree polynomial $n_{p,u,2}$ (viewed as a polynomial in β_k).

Using the quadratic formula, we verify that $n_{p,u,2}$ (see (B.104)) has three real roots and that $f_{p,u}(N, \phi_{kk})$ is linear in ϕ_{kk} and can thus be expressed as $f_{p,u}(N, \phi_{kk}) = f_{p,u}(N)\,\phi_{kk}$ where

$$f_{p,u}(N) := \frac{1}{2}\left(\sqrt{\frac{N-2}{N}} + 1\right).$$

We now show that the limits in (3.27) hold. From (B.59) and (B.65), the function $M_k(z_b, z_s, 0, \varphi_2; \psi) + u_k^0$ is strictly decreasing w.r.t. z_k and it approaches $\pm\infty$ as z_k approaches $\mp\infty$. Thus, if z_k^* is the unique solution to $M_k(z_b^*, z_s^*, 0, \varphi_2; \psi) + u_k^0 = u_k^0$, as $u_k^0 \to \infty$, $M_k(z_b^*, z_s^*, 0, \varphi_2; \psi) \to \infty$ and therefore, $z_k^* \to -\infty$. It follows that $\lim_{u_k^0 \to \infty} z_k^* = -\infty$. Similarly, $\lim_{u_k^0 \to -\infty} z_k^* = \infty$. From (B.41) and (B.58), it follows that

$$p_k^* = \frac{\left(\beta_k + e^{z_k^*}(\beta_k N - \phi_{kk})\right)\left(\beta_k\left(Ne^{z_k^*}+1\right)^2 - e^{z_k^*}\phi_{kk}\right)}{\left(1+Ne^{z_k^*}\right)\left(\beta_k\left(1+(N-1)e^{z_k^*}\right)\left(1+Ne^{z_k^*}\right) - e^{z_k^*}\phi_{kk}\right)}. \tag{B.106}$$

Taking limits in the above expression yields,

$$\lim_{u_k^0 \to -\infty} p_k^* = \frac{N}{N-1}\beta_k - \frac{\phi_{kk}}{N-1}, \text{ and}$$
$$\lim_{u_k^0 \to \infty} p_k^* = \beta_k.$$

Finally, we show that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, (i) and (ii) in Proposition 14 hold. From (B.40), (B.41) and (B.55), $\partial p_k^*/\partial u_k^0$ is a rational function w.r.t. (ϕ_{bs}, ϕ_{sb}). Moreover, at $(\phi_{bs}, \phi_{sb}) = (0, 0)$, the partial derivative $\partial p_k^*/\partial u_k^0$ is given by (B.100). Thus, $\partial p_k^*/\partial u_k^0$ is continuous w.r.t. (ϕ_{bs}, ϕ_{sb}) at $(0, 0)$. We conclude that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0) \subset \mathbb{R}^2$, cases (i) and (ii) in Proposition 14 hold true. □

Proof of Proposition 15. Using (B.106) from the Proof of Proposition 14 and (B.40), we can write

$$\pi_k^*\Big|_{\varphi_1=0} = \frac{e^{z_k^*}\left(\beta_k + e^{z_k^*}\left(\beta_k N - \phi_{kk}\right)\right)\left(\beta_k\left(Ne^{z_k^*}+1\right)^2 - e^{z_k^*}\phi_{kk}\right)}{\left(Ne^{z_k^*}+1\right)^2\left(\beta_k\left(\left(N-1\right)e^{z_k^*}+1\right)\left(Ne^{z_k^*}+1\right) - e^{z_k^*}\phi_{kk}\right)}, \quad k \in \{b, s\}. \tag{B.107}$$

It follows that $\frac{\partial \pi_k^*}{\partial u_k^0}\Big|_{\varphi_1=0} = \frac{\partial \pi_k^*}{\partial z_k^*}\frac{\partial z_k^*}{\partial u_k^0}$, where $\frac{\partial z_k^*}{\partial u_k^0}$ is given by (B.101). We show (see *Gumbel_N.nb*) that

$$\frac{\partial \pi_k^*}{\partial u_k^0}\Big|_{\varphi_1=0} = -\frac{n_{\pi,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)}{d_{\pi,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)}, \tag{B.108}$$

where $n_{\pi,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ and $d_{\pi,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ can be written as polynomials in $e^{z_k^*}$ in the following way:

$$\begin{aligned} n_{\pi,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right) &:= \sum_{m=1}^{6} n_{\pi,u,m} e^{m z_k^*}, \quad \text{and} \\ d_{\pi,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right) &:= \sum_{m=0}^{7} d_{\pi,u,m} e^{m z_k^*}. \end{aligned} \tag{B.109}$$

Moreover, the coefficients $n_{\pi,u,m}$ are as follows:

$$\begin{aligned} n_{\pi,u,1} &= \beta_k^3, \\ n_{\pi,u,2} &= \beta_k^2 \left(5\beta_k N - 4\phi_{kk}\right), \\ n_{\pi,u,3} &= \beta_k \left(10\beta_k^2 N^2 + 2\beta_k \left(1-7N\right)\phi_{kk} + 5\phi_{kk}^2\right), \\ n_{\pi,u,4} &= 10\beta_k^3 N^3 + 2N\beta_k^2 \left(2-9N\right)\phi_{kk} + \beta_k \left(9N-2\right)\phi_{kk}^2 - 2\phi_{kk}^3, \\ n_{\pi,u,5} &= \beta_k N \left(5\beta_k^2 N^3 + 2N\beta_k \left(1-5N\right)\phi_{kk} + \left(4N-1\right)\phi_{kk}^2\right) \text{ and} \\ n_{\pi,u,6} &= \beta_k N^2 \left(\beta_k^2 N^3 - 2\beta_k N^2 \phi_{kk} + \phi_{kk}^2\right). \end{aligned} \tag{B.110}$$

The coefficients $d_{\pi,u,m}$ are given by

$$d_{\pi,u,0} = \beta_k^3,$$

$$d_{\pi,u,1} = \beta_k^2 \left(\beta_k \left(7N-1\right) - 4\phi_{kk}\right),$$

$$d_{\pi,u,2} = \beta_k \left(\beta_k^2 \left(21N^2 - 7N + 1\right) + 4\beta_k \left(1-5N\right)\phi_{kk} + 5\phi_{kk}^2\right),$$

$$d_{\pi,u,3} = N\beta_k^3 \left(35N^2 - 20N + 5\right) + \beta_k^2 \left(-40N^2 + 15N - 1\right)\phi_{kk} + \beta_k \left(15N-3\right)\phi_{kk}^2 - 2\phi_{kk}^3,$$

$$d_{\pi,u,4} = N^2\beta_k^3 \left(35N^2 - 30N + 10\right) + N\beta_k^2 \left(-40N^2 + 21N - 3\right)\phi_{kk} + 5N\beta_k \left(3N-1\right)\phi_{kk}^2 - 2N\phi_{kk}^3,$$

$$d_{\pi,u,5} = \beta_k N^2 \left(N\beta_k^2 \left(21N^2 - 25N + 10\right) + \beta_k \left(-20N^2 + 13N - 3\right)\phi_{kk} + \left(5N-1\right)\phi_{kk}^2\right),$$

$$d_{\pi,u,6} = \beta_k N^3 \left(N\beta_k^2 \left(7N^2 - 11N + 5\right) + \beta_k \left(-4N^2 + 3N - 1\right)\phi_{kk} + \phi_{kk}^2\right), \text{ and}$$

$$d_{\pi,u,7} = \beta_k^3 \left(N-1\right)^2 N^5. \tag{B.111}$$

Because the expressions determining $n_{\pi,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ and $d_{\pi,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ are complex, we focus on finding sufficient conditions for these expressions to have a specific sign for all z_k^*. We verify in the supplementary file *Gumbel_N.nb* that $n_{\pi,u}$ and $d_{\pi,u}$ (see (B.109)) are positive, if either of the two conditions below, (i-a) or (i-b), hold:

(i-a) $\phi_{kk} \leq 0$, $N \geq 2$ and $\beta_k > 0$.

(i-b) $\phi_{kk} > 0$, $N \geq 2$ and $\beta_k > g_{\pi,u}(N, \phi_{kk})$, where $g_{\pi,u}(N, \phi_{kk})$ is the largest real root of the third degree polynomial $n_{\pi,u,6}$ (viewed as a polynomial in β_k).

Using the quadratic formula, we verify that $n_{\pi,u,6}$ (see (B.110)) has three real roots and that $g_{\pi,u}(N, \phi_{kk})$ is linear in ϕ_{kk} and can thus be expressed as $g_{\pi,u}(N, \phi_{kk}) = g_{\pi,u}(N)\,\phi_{kk}$ where

$$g_{\pi,u}(N) := \sqrt{\frac{N-1}{N^3}} + \frac{1}{N}.$$

We now show that the limits in (3.29) hold. From the Proof of Proposition 15, we have that $\lim_{u_k^0 \to \infty} z_k^* = -\infty$ and $\lim_{u_k^0 \to -\infty} z_k^* = \infty$ (see the paragraph above (B.106)). Taking limits in (B.107) yields,

$$\lim_{u_k^0 \to -\infty} \pi_k^* = \frac{1}{N-1}\beta_k - \frac{\phi_{kk}}{(N-1)N}, \text{ and}$$
$$\lim_{u_k^0 \to \infty} \pi_k^* = 0.$$

Finally, we show that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, the first part in Proposition 15 holds. From (B.40), (B.41) and (B.55), $\partial \pi_k^* / \partial u_k^0$ is a rational function w.r.t. (ϕ_{bs}, ϕ_{sb}). Moreover, at $(\phi_{bs}, \phi_{sb}) = (0, 0)$, the partial derivative $\partial \pi_k^* / \partial u_k^0$ is given by (B.108). Thus, $\partial \pi_k^* / \partial u_k^0$ is continuous w.r.t. (ϕ_{bs}, ϕ_{sb}) at $(0, 0)$. We conclude that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0) \subset \mathbb{R}^2$, the first part of Proposition 15 holds true. □

Proof of Proposition 16. Plugging (B.106) from the Proof of Proposition 14 and (B.40) into (3.30), we obtain

$$CS_k^*\Big|_{\varphi_1=0} = \frac{-\beta_k^2\left(1+Ne^{z_k^*}\right)^3 + \beta_k\phi_{kk}e^{z_k^*}\left(\left(2N-1\right)e^{z_k^*}+3\right)\left(1+Ne^{z_k^*}\right) - 2e^{2z_k^*}\phi_{kk}^2}{\left(1+Ne^{z_k^*}\right)\left(\beta_k\left(1+\left(N-1\right)e^{z_k^*}\right)\left(1+Ne^{z_k^*}\right) - e^{z_k^*}\phi_{kk}\right)} + \mu_k + \beta_k\left(\ln\left(N+1\right)+\gamma\right), \quad k \in \{b, s\}. \tag{B.112}$$

It follows that $\frac{\partial CS_k^*}{\partial u_k^0}\Big|_{\varphi_1=0} = \frac{\partial CS_k^*}{\partial z_k^*}\frac{\partial z_k^*}{\partial u_k^0}$, where $\frac{\partial z_k^*}{\partial u_k^0}$ is given by (B.101). We show (see *Gumbel_N.nb*) that

$$\frac{\partial CS_k^*}{\partial u_k^0}\Big|_{\varphi_1=0} = \frac{n_{CS,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)}{d_{CS,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)}, \tag{B.113}$$

where $n_{CS,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ and $d_{CS,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ can be written as polynomials in $e^{z_k^*}$

in the following way,

$$n_{CS,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right) := \sum_{m=1}^{5} n_{CS,u,m} e^{m z_k^*}, \quad \text{and}$$
$$d_{CS,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right) := \sum_{m=0}^{6} d_{CS,u,m} e^{m z_k^*}. \tag{B.114}$$

Moreover, the coefficients $n_{CS,u,m}$ are as follows:

$$\begin{aligned} n_{CS,u,1} &= \beta_k^2 \left(\beta_k - 2\phi_{kk}\right), \\ n_{CS,u,2} &= 2\beta_k \left(2\beta_k^2 N + \beta_k \left(1-4N\right)\phi_{kk} + 2\phi_{kk}^2\right), \\ n_{CS,u,3} &= 6\beta_k^3 N^2 + \beta_k^2 \left(-12N^2 + 5N - 1\right)\phi_{kk} + \beta_k \left(8N-3\right)\phi_{kk}^2 - 2\phi_{kk}^3, \\ n_{CS,u,4} &= 2\beta_k N \left(2\beta_k^2 N^2 + \beta_k \left(-4N^2 + 2N - 1\right)\phi_{kk} + \left(2N-1\right)\phi_{kk}^2\right) \text{ and} \\ n_{CS,u,5} &= \beta_k N^2 \left(\beta_k^2 N^2 + \beta_k \left(-2N^2 + N - 1\right)\phi_{kk} + \phi_{kk}^2\right). \end{aligned} \tag{B.115}$$

The coefficients $d_{CS,u,m} = d_{p,u,m}$ for all $m \in \{0, \ldots, 6\}$ where $d_{p,u,m}$ is given by (B.105). Because the expressions determining $n_{CS,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ and $d_{CS,u}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ are complex, we focus on finding sufficient conditions for these expressions to have a specific sign for all z_k^*.

Case (i): $\left.\frac{\partial CS_k^*}{\partial u_k^0}\right|_{\varphi_1=0} > 0$. We verify in the supplementary file *Gumbel_N.nb* that $n_{CS,u}$ and $d_{CS,u}$ (see (B.114)) are positive, if either of the two conditions below, (i-a) or (i-b), hold:

(i-a) $\phi_{kk} \leq 0$, $N \geq 2$ and $\beta_k > 0$.

(i-b) $\phi_{kk} > 0$, $N \geq 2$ and $\beta_k > 2\phi_{kk}$.

Case (ii): $\left.\frac{\partial CS_k^*}{\partial u_k^0}\right|_{\varphi_1=0} < 0$. We verify in the supplementary file *Gumbel_N.nb* that $n_{CS,u}$ and $d_{CS,u}$ (see (B.114)) are negative and positive, respectively, if the condition below, (ii-a), holds:

(ii-a) $\phi_{kk} > 0$, $N \geq 2$ and $f\left(N\right)\phi_{kk} < \beta_k < f_{CS,u}\left(N, \phi_{kk}\right)$, where $f_{CS,u}\left(N, \phi_{kk}\right)$ is the unique real root of the third degree polynomial $n_{CS,u,3}$ (viewed as a polynomial in β_k).

We next verify that $n_{CS,u,3}$ indeed has a unique real root and that $f_{CS,u}\left(N, \phi_{kk}\right)$ is linear in ϕ_{kk} and can thus be expressed as $f_{CS,u}\left(N, \phi_{kk}\right) = f_{CS,u}\left(N\right)\phi_{kk}$. We note that

$n_{CS,u,3}/(6N^2)$ has the standard form

$$\frac{n_{CS,u,3}}{6N^2} = \beta_k^3 + b_2\beta_k^2 + b_1\beta_k + b_0, \tag{B.116}$$

where b_0, b_1 and b_2 depend on N and ϕ_{kk}. We first clarify why this polynomial has a unique real root using Cardano's condition. We define $t_k := \frac{1}{3}(3b_1 - b_2^2)$, $s_k := \frac{1}{27}(2b_2^3 - 9b_2b_1 + 27b_0)$ and $\Delta_k := (s_k/2)^2 + (t_k/3)^3$. Equivalently,

$$\begin{aligned} t_k &= -\frac{(144N^4 - 264N^3 + 103N^2 - 10N + 1)\,\phi_{kk}^2}{108N^4}, \\ s_k &= -\frac{(1728N^6 - 4752N^5 + 4356N^4 - 1106N^3 + 192N^2 - 15N + 1)\,\phi_{kk}^3}{2916N^6}, \text{ and} \\ \Delta_k &:= \frac{(4608N^6 - 11136N^5 + 6944N^4 + 408N^3 - 97N^2 + 18N - 1)\,\phi_{kk}^6}{139968N^8}. \end{aligned} \tag{B.117}$$

For each $k \in \{b, s\}$, if $\phi_{kk} > 0$ and $N \geq 2$, then $\Delta_k > 0$. It thus follows that $n_{CS,u,3}$ has a unique real root given by

$$\alpha := \text{Car}(s_k, \Delta_k) - \frac{b_2}{3}, \tag{B.118}$$

where $b_2 = (-12N^2 + 5N - 1)\,\phi_{kk}/(6N^2)$ and $\text{Car}(\cdot,\cdot)$ is given

$$\text{Car}(s_k, \Delta_k) := \left(-\frac{s_k}{2} + \sqrt{\Delta_k}\right)^{1/3} + \left(-\frac{s_k}{2} - \sqrt{\Delta_k}\right)^{1/3}. \tag{B.119}$$

For $\phi_{kk} > 0$, it is not difficult to see that α is linear in ϕ_{kk}, so we can write

$$f_{CS,u}(N) = \alpha/\phi_{kk}, \tag{B.120}$$

where α is given by (B.118).

Finally, we show that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, (i) and (ii) in Proposition 16 hold. From (B.40), (B.41) and (B.55), $\partial CS_k^*/\partial u_k^0$ is a rational function w.r.t. (ϕ_{bs}, ϕ_{sb}). Moreover, at $(\phi_{bs}, \phi_{sb}) = (0, 0)$, the partial derivative $\partial CS_k^*/\partial u_k^0$ is given by (B.113). Thus, $\partial CS_k^*/\partial u_k^0$ is continuous w.r.t. (ϕ_{bs}, ϕ_{sb}) at $(0, 0)$. We conclude that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0) \subset \mathbb{R}^2$, cases (i) and (ii) in Proposition 16 hold true. □

Proof of Proposition 17. Suppose that $N \geq 2$ and for each $k \in \{b, s\}$, (ϕ_{kk}, β_k) satisfies (3.23). Assume that $\varphi_1 = (\phi_{bs}, \phi_{sb}) = 0$. From the proof of Proposition 12, M_k^{C} does not depend on z_j for $j \neq k$ (see (B.86)). Moreover,

$$\lim_{z_k \to -\infty} M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi) = \infty, \tag{B.121}$$

and $M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi)$ is strictly decreasing in z_k for all $z_k \in \mathbb{R}$. From (B.86), we compute

$$M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi)\Big|_{z_k=0} = \frac{2\phi_{kk}}{N+1} - \beta_k(1+N) - u_k^0. \tag{B.122}$$

The only root of (B.122) as a polynomial in β_k is given by

$$\gamma^{\mathrm{C}}(N, \phi_{kk}, u_k^0) := \frac{2\phi_{kk} - u_k^0(N+1)}{(N+1)^2}. \tag{B.123}$$

Observation (i): The condition $\beta_k > \gamma^{\mathrm{C}}(N, \phi_{kk}, u_k^0)$ implies that (B.122) is strictly negative. We combine this fact with (B.121) and the facts that $M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi)$ is strictly decreasing in z_k $\forall z_k \in \mathbb{R}$ and z_k^{C} is the unique solution of $M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi) = 0$ to conclude that $z_k^{\mathrm{C}} < 0$.

Observation (ii): If $\beta_k < \gamma^{\mathrm{C}}(N, \phi_{kk}, u_k^0)$, then (B.122) is strictly positive and thus $z_k^{\mathrm{C}} > 0$.

The Proof of Proposition 12 implies that there is $\epsilon > 0$ and a unique continuous function

$$\left(z_b^{\mathrm{C}}(\cdot), z_s^{\mathrm{C}}(\cdot)\right) : B_\epsilon(0,0) \longrightarrow \mathbb{R}^2$$

such that for all $\varphi_1 \in B_\epsilon(0,0)$, $\boldsymbol{M}^{\mathrm{C}}\left(z_b^{\mathrm{C}}(\varphi_1), z_s^{\mathrm{C}}(\varphi_1), \varphi_1, \varphi_2; \psi\right) = 0$. By this continuity, for $\hat{\varepsilon} = |z_k^{\mathrm{C}}(0,0)|/2$, there exists $\delta > 0$ such that for all $\varphi_1 \in B_{\min(\delta,\varepsilon)}(0,0)$,

$$z_k^{\mathrm{C}}(0,0) - \frac{\left|z_k^{\mathrm{C}}(0,0)\right|}{2} < z_k^{\mathrm{C}}(\varphi_1) < z_k^{\mathrm{C}}(0,0) + \frac{\left|z_k^{\mathrm{C}}(0,0)\right|}{2}.$$

Thus, when $z_k^{\mathrm{C}}(0,0) < 0$, by observation (i) we obtain $z_k^{\mathrm{C}}(\varphi_1) < \frac{z_k^{\mathrm{C}}(0,0)}{2} < 0$. When $z_k^{\mathrm{C}} > 0$, by observation (ii) we obtain $z_k^{\mathrm{C}}(\varphi_1) > \frac{z_k^{\mathrm{C}}(0,0)}{2} > 0$. Concluding the proof of the Proposition.

□

Proof of Corollary 3. Note that from (B.95) and (B.123),

$$\gamma\left(N,\phi_{kk},u_k^0\right)-\gamma^{\mathrm{C}}\left(N,\phi_{kk},u_k^0\right) = \frac{2(N-1)\phi_{kk}+\left(-N^2+N+2\right)u_k^0+\sqrt{N^2(N+1)^2\left(u_k^0\right)^2+4(N^2-1)\phi_{kk}^2-4(N-1)(N+1)^2u_k^0\phi_{kk}}}{2(N+1)^2}. \quad \text{(B.124)}$$

From (B.96) and the Proof of Corollary 1, if $\gamma\left(N,\phi_{kk},u_k^0\right)\geq 0$ then,

$$\text{either } (\phi_{kk}<0 \text{ and } u_k^0 \leq \frac{2\phi_{kk}}{N+1}) \text{ or } \phi_{kk}\geq 0.$$

Suppose that $(\phi_{kk}<0 \text{ and } u_k^0\leq \frac{2\phi_{kk}}{N+1})$. By (B.124), $\gamma\left(N,\phi_{kk},u_k^0\right)\geq\gamma^{\mathrm{C}}\left(N,\phi_{kk},u_k^0\right)$. On the other hand, if $\phi_{kk}\geq 0$ then (B.124) implies that $\gamma\left(N,\phi_{kk},u_k^0\right)\geq\gamma^{\mathrm{C}}\left(N,\phi_{kk},u_k^0\right)$.

□

Proof of Corollary 4. First assume $\phi_{kk}\leq 0$ and $\beta_k>0$. From the definition of the quantity $\gamma^{\mathrm{C}}(N,\phi_{kk},u_k^0)$ in (B.123), if $u_k^0\geq 2\phi_{kk}/(N+1)$, then $\gamma^{\mathrm{C}}(N,\phi_{kk},u_k^0)\leq 0$. In this case, statement (ii) of Proposition 17 is not feasible as $\beta_k>0$.

Now assume that $\phi_{kk}>0$ and $\beta_k>\frac{8\phi_{kk}}{27N}$. The unique $\tilde{u}_k^{\mathrm{C}}$ such that $\gamma^{\mathrm{C}}(N,\phi_{kk},\tilde{u}_k^{\mathrm{C}})=\frac{8\phi_{kk}}{27N}$ is given by

$$\tilde{u}_k^0=-\frac{2\left(N\left(4N-19\right)+4\right)\phi_{kk}}{27N\left(N+1\right)}.$$

If $u_k^0\geq\tilde{u}_k^{\mathrm{C}}$, then $\gamma^{\mathrm{C}}(N,\phi_{kk},\tilde{u}_k^0)\leq\frac{8\phi_{kk}}{27N}$. Thus, statement (ii) in Proposition 17 is not feasible as $\beta_k>\frac{8\phi_{kk}}{27N}$.

We finish the proof with a piecewise definition for $\tilde{u}_k^{\mathrm{C}}$,

$$\tilde{u}_k^{\mathrm{C}}:=\begin{cases}\frac{2\phi_{kk}}{N+1} & \text{if } \phi_{kk}\leq 0,\\ -\frac{2(N(4N-19)+4)\phi_{kk}}{27N(N+1)} & \text{if } \phi_{kk}>0.\end{cases} \quad \text{(B.125)}$$

□

Proof of Proposition 18. First, we set $\varphi_1=(\phi_{bs},\phi_{sb})=0$. From (B.58) and (B.86), we obtain

$$M_k(z_b, z_s, 0, \varphi_2; \psi) - M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi) = \frac{\beta_k (N-1) e^{z_k} \left(\beta_k (Ne^{z_k}+1)^2 - e^{z_k}\phi_{kk}\right)}{\beta_k ((N-1)e^{z_k}+1)(Ne^{z_k}+1) - e^{z_k}\phi_{kk}} \tag{B.126}$$

Step (i): If $\phi_{kk} \leq 0$ or ($\phi_{kk} > 0$ and $\beta_k > f(N)\phi_{kk}$), then (B.126) is strictly positive for all $z_k \in \mathbb{R}$. The proofs of Propositions 10 and 12 imply that the functions $M_k(z_b, z_s, 0, \varphi_2; \psi)$ and $M_k^{\mathrm{C}}(z_b, z_s, 0, \varphi_2; \psi)$ are independent of z_l for $l \neq k$ and are strictly decreasing on z_k. It follows that $z_k^* > z_k^{\mathrm{C}}$.

Again, from the Proofs of Proposition 10 and 12, there is $\epsilon > 0$ and unique continuous functions $(z_b^*(\cdot), z_s^*(\cdot)) : B_\epsilon(0,0) \longrightarrow \mathbb{R}^2$ and $\left(z_b^{\mathrm{C}}(\cdot), z_s^{\mathrm{C}}(\cdot)\right) : B_\epsilon(0,0) \longrightarrow \mathbb{R}^2$ such that for all $\varphi_1 \in B_\epsilon(0,0)$,

$$M(z_b^*(\varphi_1), z_s^*(\varphi_1), \varphi_1, \varphi_2; \psi) = 0 = M^{\mathrm{C}}\left(z_b^{\mathrm{C}}(\varphi_1), z_s^{\mathrm{C}}(\varphi_1), \varphi_1, \varphi_2; \psi\right). \tag{B.127}$$

By continuity, there is $\hat{\epsilon} > 0$ such that $z_k^*(\varphi_1) > z_k^{\mathrm{C}}(\varphi_1)$ for all $\varphi_1 \in B_{\hat{\epsilon}}(0,0)$.

Step (ii): By (3.17) and (3.22), $x_k^* = \omega(z_k^*)$ and $x_k^{\mathrm{C}} = \omega\left(z_k^{\mathrm{C}}\right)$ where $\omega(z) = \frac{1}{e^{-z}+N}$. Note that $\omega'(z) > 0$ for all $z \in \mathbb{R}$. Then, $x_k^* = \omega(z_k^*) > \omega\left(z_k^{\mathrm{C}}\right) = x_k^{\mathrm{C}}$ for all $\varphi_1 \in B_{\hat{\epsilon}}(0,0)$.

Step (iii): From (3.12), the function $\boldsymbol{p}(\boldsymbol{z}) = \Phi\Omega(\boldsymbol{z}) - \boldsymbol{\beta}\boldsymbol{z} - \boldsymbol{u}_0$ determines the equilibrium prices p_k^* and p_k^{C}. Note that

$$\frac{\partial p_k}{\partial z_k} = \phi_{kk}\omega'(z_k) - \beta_k = \frac{\phi_{kk}e^{z_k} - \beta_k (Ne^{z_k}+1)^2}{(Ne^{z_k}+1)^2}. \tag{B.128}$$

The function $z_k \mapsto \frac{e^{z_k}}{(Ne^{z_k}+1)^2}$ has a unique maximum over $\mathbb{R}$ at $z_k^0 = \log\left(\frac{1}{N}\right)$ and such maximum is given by $\frac{1}{4N}$. Then, (B.128) is strictly negative for all $z_k \in \mathbb{R}$ when either $\phi_{kk} \leq 0$ or ($\phi_{kk} > 0$ and $\beta_k > \frac{\phi_{kk}}{4N}$). Moreover, from (3.18), $f(N)\phi_{kk} > \frac{\phi_{kk}}{4N}$ for all $N \geq 2$ and $\phi_{kk} > 0$. Then, for $\varphi_1 = 0$, $p_k^* = p_k(z_k^*) < p_k\left(z_k^{\mathrm{C}}\right) = p_k^{\mathrm{C}}$. By step (i) and continuity, there is $\hat{\epsilon} > 0$ such that $p_k^*(z_k^*(\varphi_1)) < p_k^{\mathrm{C}}\left(z_k^C(\varphi_1)\right)$ for all $\varphi_1 \in B_{\hat{\epsilon}}(0,0)$.

□

Proof of Proposition 19. Recall from (B.41) in the Proof of Proposition 9,

$$\boldsymbol{p}^* = \Phi\Omega(\boldsymbol{z}^*) - \boldsymbol{\beta}\boldsymbol{z}^* - \boldsymbol{u}_0,$$

where $\Omega(\boldsymbol{z}^*) = (\omega(z_b^*), \omega(z_s^*))^T$, $\omega(\cdot)$ and $\boldsymbol{z}^*$ are given by (B.40) and (B.55), respectively. We want to compute the following quantity when $\varphi_1 = 0$,

$$\frac{\partial p_k^*}{\partial N} = \left(\phi_{kk}\omega'\left(z_k^*\right) - \beta_k\right)\frac{\partial z_k^*}{\partial N} + \phi_{kk}\frac{\partial \omega\left(z_k^*\right)}{\partial N} + \phi_{kj}\left(\omega'(z_j^*)\frac{\partial z_j^*}{\partial N} + \frac{\partial \omega(z_j^*)}{\partial N}\right), \quad \text{(B.129)}$$

for $k, j \in \{b, s\}$ $k \neq j$. By (B.40), $\omega(z_k^*) = 1/(e^{-z_k^*} + N)$, then

$$\begin{aligned} \omega'\left(z_k^*\right) &= \frac{e^{-z_k^*}}{\left(e^{-z_k^*} + N\right)^2} \quad \text{and} \\ \frac{\partial \omega\left(z_k^*\right)}{\partial N} &= \frac{-1}{\left(e^{-z_k^*} + N\right)^2}. \end{aligned} \quad \text{(B.130)}$$

After differentiating both sides of the two equations in (B.55) w.r.t. N, we can write $\partial z_k^*/\partial N$ at $\varphi_1 = (\phi_{bs}, \phi_{sb}) = 0$,

$$\left[\begin{array}{c} \frac{\partial z_s^*}{\partial N} \\ \frac{\partial z_b^*}{\partial N} \end{array}\right]_{\varphi_1=0} = \frac{1}{J_N}\left[\begin{array}{cc} -d & b \\ c & -a \end{array}\right]\left[\begin{array}{c} \frac{\partial[h_b - \phi_{ss}d_bL_b]}{\partial N}\omega\left(z_b^*\right) + \left(2\phi_{bb} - \phi_{ss}d_bL_b + h_b\right)\frac{\partial\omega(z_b^*)}{\partial N} \\ \frac{\partial[h_s - \phi_{bb}d_sL_s]}{\partial N}\omega\left(z_s^*\right) + \left(2\phi_{ss} - \phi_{bb}d_sL_s + h_s\right)\frac{\partial\omega(z_s^*)}{\partial N} \end{array}\right], \quad \text{(B.131)}$$

where $J_N := ad - bc$ and

$$\begin{aligned} a &= \omega\left(z_b^*\right)\left(\frac{\partial h_b}{\partial z_s} - \phi_{ss}\frac{\partial\left[d_bL_b\right]}{\partial z_s}\right), \\ b &= \omega\left(z_b^*\right)\left(\frac{\partial h_b}{\partial z_b} - \phi_{ss}\frac{\partial\left[d_bL_b\right]}{\partial z_b}\right) + \left(2\phi_{bb} - \phi_{ss}d_bL_b + h_b\right)\omega'\left(z_b^*\right) - \beta_b, \\ c &= \omega\left(z_s^*\right)\frac{\partial\left[h_s - \phi_{bb}d_sL_s\right]}{\partial z_s} + \left(2\phi_{ss} - \phi_{bb}d_sL_s + h_s\right)\omega'\left(z_s^*\right) - \beta_s, \quad \text{and} \\ d &= \omega\left(z_s^*\right)\frac{\partial\left[h_s - \phi_{bb}d_sL_s\right]}{\partial z_b}. \end{aligned} \quad \text{(B.132)}$$

Moreover, recall that L_k, d_k and h_k for $k \in \{b, s\}$ are given by (3.15). After plugging (B.132) into (B.131) and then plugging the resulting expression into (B.129), we show (see *Gumbel_N.nb*) that

$$\left.\frac{\partial p_k^*}{\partial N}\right|_{\varphi_1=0} = \frac{n_p\left(z_k^*, \beta_k, \phi_{kk}, N\right)}{d\left(z_k^*, \beta_k, \phi_{kk}, N\right)}, \quad \text{(B.133)}$$

where $n_p\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ and $d\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ can be written as polynomials in $e^{z_k^*}$ in the following way,

$$\begin{aligned} n_p\left(z_k^*, \beta_k, \phi_{kk}, N\right) &:= \sum_{m=2}^{6} n_{m,p} e^{m z_k^*}, \quad \text{and} \\ d\left(z_k^*, \beta_k, \phi_{kk}, N\right) &:= \sum_{m=0}^{6} d_m e^{m z_k^*}. \end{aligned} \tag{B.134}$$

Moreover, the coefficients $n_{m,p}$ are as follows:

$$\begin{aligned} n_{2,p} &= \beta_k^3\left(\phi_{kk} - \beta_k\right), \\ n_{3,p} &= -\beta_k^2\left(4\beta_k^2 N + \beta_k\phi_{kk}\left(1-4N\right) + 2\phi_{kk}^2\right), \\ n_{4,p} &= \beta_k\left(-6\beta_k^3 N^2 + 2\beta_k^2\phi_{kk}N\left(3N-1\right) + \beta_k\phi_{kk}^2\left(3-4N\right) + \phi_{kk}^3\right), \\ n_{5,p} &= -\beta_k\left(4\beta_k^3 N^3 + \beta_k^2\phi_{kk}N^2\left(1-4N\right) + \beta_k\phi_{kk}^2 N\left(2N-3\right) + \phi_{kk}^3\right), \text{ and} \\ n_{6,p} &= \beta_k^3 N^4\left(\phi_{kk} - \beta_k\right). \end{aligned} \tag{B.135}$$

The coefficients d_m are given by

$$d_0 = \beta_k^3,$$

$$d_1 = \beta_k^2\left(\beta_k\left(6N-1\right) - 4\phi_{kk}\right),$$

$$d_2 = \beta_k\left(\beta_k^2\left(15N^2 - 6N + 1\right) + 4\beta_k\left(1-4N\right)\phi_{kk} + 5\phi_{kk}^2\right),$$

$$d_3 = \beta_k^3 N\left(20N^2 - 14N + 4\right) + \beta_k^2\phi_{kk}\left(-24N^2 + 11N - 1\right) + \beta_k\phi_{kk}^2\left(10N-3\right) - 2\phi_{kk}^3,$$

$$d_4 = \beta_k N\left(\beta_k^2 N\left(15N^2 - 16N + 6\right) + \beta_k\phi_{kk}\left(-16N^2 + 10N - 2\right) + \left(5N-2\right)\phi_{kk}^2\right),$$

$$d_5 = \beta_k N^2\left(\beta_k^2 N\left(6N^2 - 9N + 4\right) + \beta_k\phi_{kk}\left(-4N^2 + 3N - 1\right) + \phi_{kk}^2\right), \text{ and}$$

$$d_6 = \beta_k^3\left(N-1\right)^2 N^4. \tag{B.136}$$

Because the expressions determining $n_p\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ and $d\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ are complex, we focus on finding sufficient conditions for these expressions to have a specific sign for all z_k^*. Note that not all the coefficients in $\{d_m\}_{m=0}^{6}$ can be simultaneously negative because $d_0 > 0$. Thus, we focus on finding the regions for which $d_m > 0$ for each

$m = 0, \ldots, 6$.

Case (i): $\frac{\partial p_k^*}{\partial N}\Big|_{\varphi_1=0} < 0$. We verify in the supplementary file *Gumbel_N.nb* that n_p and d (see (B.134)) are negative and positive, respectively, if either of the two conditions below, (i-a) or (i-b), hold:

(i-a) $\phi_{kk} > 0$, $N \geq 2$ and $\beta_k > \phi_{kk}$.

(i-b) $\phi_{kk} \leq 0$, $N \geq 2$ and $\beta_k > g_p(N, \phi_{kk})$, where $g_p(N, \phi_{kk})$ is the largest real root of the third degree polynomial $n_{5,p}/\beta_k$ (viewed as a polynomial in β_k).

We next verify that $n_{5,p}/\beta_k$ has three real roots and that $g_p(N, \phi_{kk})$ is linear in ϕ_{kk} and can thus be expressed as $g_p(N, \phi_{kk}) = g_p(N)\,\phi_{kk}$. We note that $n_{5,p}/(\beta_k(4N^3))$ has the standard form

$$\frac{n_{5,p}}{\beta_k(4N^3)} = \beta_k^3 + c_2\beta_k^2 + c_1\beta_k + c_0, \tag{B.137}$$

where c_0, c_1 and c_2 depend on N and ϕ_{kk}. Following the proof of Proposition 10, we use Cardano's formula to define

$$\begin{aligned} t_k &:= \frac{1}{3}(3c_1 - c_2^2) = -\frac{(16N^2 - 32N + 37)\,\phi_{kk}^2}{48N^2}, \\ s_k &:= \frac{1}{27}(2c_2^3 - 9c_2c_1 + 27c_0) = -\frac{(64N^3 - 192N^2 + 264N - 271)\,\phi_{kk}^3}{864N^3}, \text{ and} \\ \Delta_k &:= (s_k/2)^2 + (t_k/3)^3 = -\frac{(64N^4 - 96N^3 + 52N^2 + 108N - 211)\,\phi_{kk}^6}{27648N^6}. \end{aligned} \tag{B.138}$$

For each $k \in \{b, s\}$, if $\phi_{kk} \neq 0$ and $N \geq 2$, then $\Delta_k < 0$. It thus follows that $n_{5,p}/\beta_k$ has three simple real roots given by

$$\begin{aligned} \alpha_j &:= 2\sqrt{-\frac{t_k}{3}}\cos\left(\frac{\theta + 2j\pi}{3}\right) - \frac{c_2}{3}, \quad \text{for } j = 0, 1, 2, \\ &\qquad 0 < \theta < \pi, \text{ and } \cos(\theta) = \frac{-s_k/2}{\sqrt{-(t_k/3)^3}}. \end{aligned} \tag{B.139}$$

From (B.138) and (B.139), $\theta = \theta(N)$ is a function of N and is independent of ϕ_{kk}. Moreover, for $\phi_{kk} < 0$, $\sqrt{-t_k/3}$ and $c_2 = -\frac{(4N-1)\phi_{kk}}{4N}$ are linear functions of ϕ_{kk}. Thus, for

$\phi_{kk} < 0$, the three simple real roots of $n_{5,p}/\beta_k$ can be written as

$$\alpha_j = w_j(N)\phi_{kk}. \tag{B.140}$$

We can thus write $g_p(N, \phi_{kk}) = g_p(N)\,\phi_{kk}$, where

$$g_p(N) = \max_{j=0,1,2}\{w_j(N)\}. \tag{B.141}$$

From (B.135) and (B.137),

$$\lim_{N\to\infty} \frac{n_{5,p}}{\beta_k(4N^3)} = \beta_k^3 - \phi_{kk}\beta_k^2.$$

Thus, for $\phi_{kk} < 0$, the quantity $g_p(N)\,\phi_{kk}$ approaches 0 as $N \to \infty$.

Case (ii): $\left.\frac{\partial p_k^*}{\partial N}\right|_{\varphi_1=0} > 0$. We verify in the supplementary file *Gumbel_N.nb* that n_p and d (see (B.134)) are positive if either of the two conditions below, (ii-a) or (ii-b), hold:

(ii-a) $\phi_{kk} > 0$, $N \geq 3$ and $f(3)\phi_{kk} < \beta_k < \frac{2}{3}\phi_{kk}$, where $f(3)$ is given by (3.18).

(ii-b) $\phi_{kk} > 0$, $N \geq 4$ and $f(N)\,\phi_{kk} < \beta_k < f_p(N, \phi_{kk})$, where $f_p(N, \phi_{kk})$ is the unique real root of the third degree polynomial $n_{4,p}/\beta_k$ (viewed as a polynomial in β_k).

We next verify that $n_{4,p}/\beta_k$ indeed has a unique real root and that $f_p(N, \phi_{kk})$ is linear in ϕ_{kk} and can thus be expressed as $f_p(N, \phi_{kk}) = f_p(N)\,\phi_{kk}$. We note that $n_{4,p}/(\beta_k(-6N^2))$ has the standard form

$$\frac{n_{4,p}}{\beta_k(-6N^2)} = \beta_k^3 + b_2\beta_k^2 + b_1\beta_k + b_0, \tag{B.142}$$

where b_0, b_1 and b_2 depend on N and ϕ_{kk}. We first clarify why this polynomial has a unique real root using Cardano's condition. We define

$$\begin{aligned}
\tilde{t}_k &:= -\frac{(18N^2 - 48N + 29)\,\phi_{kk}^2}{54N^2},\\
\tilde{s}_k &:= -\frac{(108N^3 - 432N^2 + 630N - 85)\,\phi_{kk}^3}{1458N^3}, \text{ and}\\
\tilde{\Delta}_k &:= \frac{(72N^4 - 168N^3 - 88N^2 + 556N - 171)\,\phi_{kk}^6}{34992N^6}.
\end{aligned} \tag{B.143}$$

For each $k \in \{b, s\}$, if $\phi_{kk} > 0$ and $N \geq 4$, then $\tilde{\Delta}_k > 0$. It thus follows that $n_{4,p}/\beta_k$ has a unique real root given by

$$\tilde{\alpha}_j := \mathrm{Car}(\tilde{s}_k, \tilde{\Delta}_k) - \frac{b_2}{3}, \tag{B.144}$$

where $b_2 = (1-3N)\phi_{kk}/(3N)$ and $\mathrm{Car}(\cdot,\cdot)$ is given

$$\mathrm{Car}(s_k, \Delta_k) := \left(-\frac{s_k}{2} + \sqrt{\Delta_k}\right)^{1/3} + \left(-\frac{s_k}{2} - \sqrt{\Delta_k}\right)^{1/3}. \tag{B.145}$$

For $\phi_{kk} > 0$, it is not difficult to see that $\tilde{\alpha}_j$ is linear in ϕ_{kk}, so we can write

$$f_p(N) = \tilde{\alpha}_j/\phi_{kk}, \tag{B.146}$$

where $\tilde{\alpha}_j$ is given by (B.144). Finally, for any $\phi_{kk} > 0$, by (B.142)

$$\lim_{N\to\infty} \frac{n_{4,p}}{\beta_k(-6N^2)} = \beta_k^2(\beta_k - \phi_{kk}).$$

It thus follows that $f_p(N)\,\phi_{kk}$ converges to ϕ_{kk} as $N \to \infty$.

From (ii-a) and (ii-b), it follows that $\frac{\partial p_k^*}{\partial N}|_{\varphi_1=0} > 0$ whenever $N \geq 3$, $\phi_{kk} > 0$ and $f(N)\,\phi_{kk} < \beta_k < f_p(N)\,\phi_{kk}$, where the definition of $f_p(N)$ is extended to include the case $N = 3$ as follows,

$$f_p(N) := \begin{cases} \frac{2}{3} & N = 3, \\ \tilde{\alpha}_j/\phi_{kk} & N \geq 4, \end{cases} \tag{B.147}$$

where $\tilde{\alpha}_j$ is given by (B.144).

Finally, we show that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, (i) and (ii) in Proposition 19 hold. From (B.129), (B.131) and (B.132), $\partial p_k^*/\partial N$ is a rational function w.r.t. (ϕ_{bs}, ϕ_{sb}). Moreover, at $(\phi_{bs}, \phi_{sb}) = (0,0)$, the partial derivative $\partial p_k^*/\partial N$ is given by (B.133). Thus, $\partial p_k^*/\partial N$ is continuous w.r.t. (ϕ_{bs}, ϕ_{sb}) at $(0,0)$. We conclude that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0) \subset \mathbb{R}^2$, cases (i) and (ii) in Proposition 19 hold true. □

Proof of Proposition 20. Using the same ideas from the Proof of Proposition 19, we want to compute the following quantity when $\varphi_1 = 0$,

$$\frac{\partial\left(Nx_k^*\right)}{\partial N} = x_k^* + N\frac{\partial x_k^*}{\partial N}. \tag{B.148}$$

By (B.40), $x_k^* = \frac{1}{e^{-z_k^*}+N}$. Then,

$$\frac{\partial x_k^*}{\partial N} = \frac{-1}{\left(e^{-z_k^*}+N\right)^2}\left(1 - e^{-z_k^*}\frac{\partial z_k^*}{\partial N}\right). \tag{B.149}$$

Using (B.131) and (B.149), we show (see *Gumbel_N.nb*) that

$$\left.\frac{\partial\left(Nx_k^*\right)}{\partial N}\right|_{\varphi_1=0} = \frac{n_{Nx}\left(z_k^*, \beta_k, \phi_{kk}, N\right)}{d_{Nx}\left(z_k^*, \beta_k, \phi_{kk}, N\right)}, \tag{B.150}$$

where

$$\begin{aligned} n_{Nx}\left(z_k^*, \beta_k, \phi_{kk}, N\right) &:= \sum_{m=1}^{6} n_{m,Nx} e^{mz_k^*} \quad \text{and} \\ d_{Nx}\left(z_k^*, \beta_k, \phi_{kk}, N\right) &:= \sum_{m=0}^{7} d_{m,Nx} e^{mz_k^*}. \end{aligned} \tag{B.151}$$

Moreover, the coefficients $n_{m,Nx}$ are as follows:

$$n_{1,Nx} = \beta_k^3,$$

$$n_{2,Nx} = \beta_k^2\left(\beta_k\left(5N-1\right) - 4\phi_{kk}\right),$$

$$n_{3,Nx} = \beta_k\left(\beta_k^2\left(10N^2-4N+1\right) + 2\beta_k\left(2-7N\right)\phi_{kk} + 5\phi_{kk}^2\right),$$

$$n_{4,Nx} = \beta_k^3 N\left(10N^2-6N+3\right) + \beta_k^2\phi_{kk}\left(-18N^2+10N-1\right) + 3\beta_k\phi_{kk}^2\left(3N-1\right) - 2\phi_{kk}^3,$$

$$n_{5,Nx} = \beta_k N\left(\beta_k^2 N\left(5N^2-4N+3\right) + 2\beta_k\phi_{kk}\left(-5N^2+4N-1\right) + \left(4N-3\right)\phi_{kk}^2\right),$$

$$n_{6,Nx} = \beta_k^2 N^2\left(\beta_k N\left(N^2-N+1\right) - \left(2N^2-2N+1\right)\phi_{kk}\right). \tag{B.152}$$

The coefficients $d_{m,Nx}$ are given by

$$d_{0,Nx} = \beta_k^3,$$

$$d_{1,Nx} = \beta_k^2 \left(\beta_k \left(7N-1\right) - 4\phi_{kk}\right),$$

$$d_{2,Nx} = \beta_k \left(\beta_k^2 \left(21N^2 - 7N + 1\right) + 4\beta_k \left(1-5N\right)\phi_{kk} + 5\phi_{kk}^2\right),$$

$$d_{3,Nx} = 5\beta_k^3 N \left(7N^2 - 4N + 1\right) + \beta_k^2 \phi_{kk} \left(-40N^2 + 15N - 1\right) + 3\beta_k \left(5N-1\right)\phi_{kk}^2 - 2\phi_{kk}^3,$$

$$d_{4,Nx} = N \left(5\beta_k^3 N \left(7N^2 - 6N + 2\right) + \beta_k^2 \phi_{kk} \left(-40N^2 + 21N - 3\right) + 5\beta_k \left(3N-1\right)\phi_{kk}^2 - 2\phi_{kk}^3\right),$$

$$d_{5,Nx} = \beta_k N^2 \left(\beta_k^2 N \left(21N^2 - 25N + 10\right) + \beta_k \left(-20N^2 + 13N - 3\right)\phi_{kk} + \left(5N-1\right)\phi_{kk}^2\right),$$

$$d_{6,Nx} = \beta_k N^3 \left(\beta_k^2 N \left(7N^2 - 11N + 5\right) + \beta_k \left(-4N^2 + 3N - 1\right)\phi_{kk} + \phi_{kk}^2\right),$$

$$d_{7,Nx} = \beta_k^3 \left(N-1\right)^2 N^5. \tag{B.153}$$

Because the expressions determining $n_{Nx}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ and $d_{Nx}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ are complex, we focus instead on finding sufficient conditions for these expressions to have a specific sign for all z_k^*. Note that not all coefficients in $\{d_{m,Nx}\}_{m=0}^{7}$ can be simultaneously negative, as $d_{0,Nx} > 0$. Thus, we focus on finding the regions for which $d_{m,Nx} > 0$ for each $m = 0, \ldots, 7$. By the previous argument, we also focus on finding regions where the coefficients $n_{m,Nx} > 0$ for all $m = 1, \ldots, 6$. We verify in the supplementary file *Gumbel_N.nb* that the coefficients $n_{m,Nx}$ and $d_{m,Nx}$ in (B.152) and (B.153), respectively, are positive if either of the two conditions below, (a) or (b), hold:

(a) $\phi_{kk} \leq 0$, $N \geq 2$ and $\beta_k > 0$.

(b) $\phi_{kk} > 0$, $N \geq 2$ and $\beta_k > g_x(N)\phi_{kk}$, where $g_x(N)\phi_{kk}$ is the unique non-zero real root of $n_{6,Nx}$ (viewed as a polynomial in β_k) and $g_x(N)$ is given by

$$g_x(N) := \frac{\left(2N^2 - 2N + 1\right)}{N\left(N^2 - N + 1\right)}. \tag{B.154}$$

We remark that finding the non-zero root of $n_{6,Nx}$ leads to a solution of a linear equation (see (B.152)), so the uniqueness of the root and its linearity in ϕ_{kk} are obvious. We also show (see *Gumbel_N.nb*) that if $N \geq 2$, $\beta_k > 0$ and $\phi_{kk} > 0$, then $g_x(N)\phi_{kk} \geq f(N)\,\phi_{kk}$.

Finally, we show that there exists ϵ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, Proposition 20 holds. Note that from (B.131) and (B.132), $\partial(Nx_k^*)/\partial N$ is a rational function w.r.t. (ϕ_{bs}, ϕ_{sb}). Moreover, at $(\phi_{bs}, \phi_{sb}) = (0,0)$, $\partial(Nx_k^*)/\partial N$ is given by (B.150). Thus, $\partial(Nx_k^*)/\partial N$ is continuous w.r.t. (ϕ_{bs}, ϕ_{sb}) at $(0,0)$. We conclude that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0) \subset \mathbb{R}^2$, $\partial(Nx_k^*)/\partial N > 0$ provided that either (a) or (b) above hold. □

Proof of Proposition 21. Using the same ideas from the Proof of Proposition 19, we want to compute the following quantity when $\varphi_1 = 0$,

$$\left.\frac{\partial CS_k^*}{\partial N}\right|_{\varphi_1=0} = \beta_k \left(\frac{1}{N+1} + \left.\frac{\partial z_k^*}{\partial N}\right|_{\varphi_1=0} \right). \tag{B.155}$$

Using equations (B.129), (B.131) and (B.149), we show (see *Gumbel_N.nb*) that

$$\left.\frac{\partial CS_k^*}{\partial N}\right|_{\varphi_1=0} = \frac{n_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right)}{d_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right)}, \tag{B.156}$$

where $n_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ and $d_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ can be written as polynomials in $e^{z_k^*}$ in the following way,

$$\begin{aligned} n_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right) &:= \sum_{m=0}^{6} n_{m,CSk} e^{m z_k^*}, \qquad \text{and} \\ d_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right) &:= \sum_{m=0}^{6} d_{m,CSk} e^{m z_k^*}. \end{aligned} \tag{B.157}$$

Moreover, the coefficients $n_{m,CSk}$ and $d_{m,CSk}$ are polynomials on $\{\phi_{kk}, \beta_k, N\}$ and are defined, respectively, as

$$n_{0,CSk} = \beta_k^4,$$

$$n_{1,CSk} = \beta_k^3 \left(\beta_k (6N-1) - 4\phi_{kk}\right),$$

$$n_{2,CSk} = \beta_k^2 \left(\beta_k^2 \left(15N^2 - 5N + 2\right) + 2\beta_k (1-9N)\,\phi_{kk} + 5\phi_{kk}^2\right),$$

$$n_{3,CSk} = \beta_k \left(\beta_k^3 N \left(20N^2 - 10N + 8\right) + \beta_k^2 \phi_{kk} \left(-32N^2 + 6N + 2\right) + \beta_k \phi_{kk}^2 (14N+1) - 2\phi_{kk}^3\right),$$

$$n_{4,CSk} = \beta_k \left(\beta_k^3 N^2 \left(15N^2 - 10N + 12\right) + \beta_k^2 \phi_{kk} \left(-28N^3 + 6N^2 + 5N - 1\right)\right)$$

$$+\beta_k \left(\beta_k \phi_{kk}^2 \left(13N^2 + 2N - 4\right) - 2\left(N+1\right)\phi_{kk}^3\right),$$

$$n_{5,CSk} = \beta_k^2 N \left(\beta_k^2 N^2 \left(6N^2 - 5N + 8\right) + \beta_k \phi_{kk} \left(-12N^3 + 2N^2 + 4N - 2\right) + \phi_{kk}^2 \left(4N^2 + N - 4\right)\right),$$

$$n_{6,CSk} = \beta_k^3 N^2 \left(\beta_k N^2 \left(N^2 - N + 2\right) + \left(N - 1 - 2N^3\right)\phi_{kk}\right). \tag{B.158}$$

The coefficients $d_{m,CSk}$ are given by

$$d_{0,CSk} = \beta_k^3 \left(N+1\right),$$

$$d_{1,CSk} = \beta_k^2 \left(N+1\right)\left(\beta_k \left(6N-1\right) - 4\phi_{kk}\right),$$

$$d_{2,CSk} = \beta_k \left(N+1\right)\left(\beta_k^2 \left(15N^2 - 6N + 1\right) + 4\beta_k \left(1 - 4N\right)\phi_{kk} + 5\phi_{kk}^2\right),$$

$$d_{3,CSk} = \left(N+1\right)\left(\beta_k^3 N \left(20N^2 - 14N + 4\right) + \beta_k^2 \phi_{kk} \left(-24N^2 + 11N - 1\right) + \beta_k \phi_{kk}^2 \left(10N - 3\right) - 2\phi_{kk}^3\right),$$

$$d_{4,CSk} = \beta_k N \left(N+1\right)\left(\beta_k^2 N \left(15N^2 - 16N + 6\right) + \beta_k \phi_{kk} \left(-16N^2 + 10N - 2\right) + \left(5N - 2\right)\phi_{kk}^2\right),$$

$$d_{5,CSk} = \beta_k N^2 \left(N+1\right)\left(\beta_k^2 N \left(6N^2 - 9N + 4\right) + \beta_k \phi_{kk} \left(-4N^2 + 3N - 1\right) + \phi_{kk}^2\right),$$

$$d_{6,CSk} = \beta_k^3 \left(N-1\right)^2 N^4 \left(N+1\right). \tag{B.159}$$

Because the expressions determining $n_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ and $d_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ are complex, we focus instead on finding sufficient conditions for these expressions to have a specific sign for all z_k^*. Note that not all coefficients in $\{d_{m,CSk}\}_{m=0}^{6}$ can be simultaneously negative, as $d_{0,CSk} > 0$. Thus, we focus on finding the regions for which $d_{m,CSk} > 0$ for each $m = 0, \ldots, 6$.

Case (i): $\left.\frac{\partial CS_k^*}{\partial N}\right|_{\varphi_1=0} > 0$. We verify in the supplementary file *Gumbel_N.nb* that $n_{m,CSk} > 0$ and $d_{m,CSk} > 0$ for all $m = 0, \ldots, 6$ (see (B.158) and (B.159)), if either of the two conditions below, (a-i) or (b-i), hold:

(a-i) $\phi_{kk} \leq 0$, $N \geq 2$ and $\beta_k > 0$.

(b-i) $\phi_{kk} > 0$, $N \geq 2$ and $\beta_k > g_{CS}(N)\phi_{kk}$, where $g_{CS}(N)\phi_{kk}$ is the unique non-zero real root of $n_{6,CSk}$ (viewed as a polynomial in β_k) and $g_{CS}(N)$ is given by

$$g_{CS}(N) := \frac{2N^3 - N + 1}{N^2\left(N^2 - N + 2\right)}. \tag{B.160}$$

We remark that finding the non-zero root of $n_{6,CSk}$ leads to a solution of a linear equation (see (B.158)), so the uniqueness of the root and its linearity in ϕ_{kk} are obvious.

Case (ii): $\left.\frac{\partial CS_k^*}{\partial N}\right|_{\varphi_1=0} < 0$. First, we want to show that for all $\phi_{kk} > 0$,

$$n_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right) = \sum_{m=0}^{6} n_{m,CSk} e^{m z_k^*} < 0.$$

Using (3.18) and (B.160), we show (see *Gumbel_N.nb*) that for all $N \geq 2$ and $\phi_{kk} > 0$,

$$f\left(N\right)\phi_{kk} < g_{CS}(N)\phi_{kk}. \tag{B.161}$$

For $N \geq 7$, $\phi_{kk} > 0$ and $\beta_k \in \left(f\left(N\right)\phi_{kk}, g_{CS}(N)\phi_{kk}\right)$, we show (see *Gumbel_N.nb*) that

$$\begin{aligned} & n_{0,CSk} > 0 \text{ for all } k = 0, \cdots 5, \text{ and} \\ & n_{m,CS6} < 0. \end{aligned} \tag{B.162}$$

From (B.162), if $N \geq 7$, $\phi_{kk} > 0$ and $\beta_k \in \left(f\left(N\right)\phi_{kk}, g_{CS}(N)\phi_{kk}\right)$, then

$$\begin{aligned} & \sum_{m=0}^{6} n_{m,CSk} e^{m z_k^*} \\ & = \underbrace{n_{0,CSk} + n_{1,CSk} e^{z_k^*} + n_{2,CSk} e^{2z_k^*} + n_{3,CSk} e^{3z_k^*} + n_{4,CSk} e^{4z_k^*} + n_{5,CSk} e^{5z_k^*}}_{>0} + \underbrace{n_{6,CSk} e^{6z_k^*}}_{<0}. \end{aligned} \tag{B.163}$$

From Proposition 13, if $\beta_k < \gamma\left(N, \phi_{kk}, u_k^0\right)$, then $z_k^* > 0$. Moreover, by hypothesis,

$z_k^* < \frac{1}{5}\ln 2$, then $e^{5z_k^*} < 2$. It follows from (B.163) that

$$\sum_{m=0}^{6} n_{m,CSk} e^{mz_k^*} < n_{0,CS_k} + 2\left(n_{1,CS_k} + n_{2,CS_k} + n_{3,CS_k} + n_{4,CS_k} + n_{5,CS_k}\right) + n_{6,CS_k}$$
$$= \beta_k \sum_{m=0}^{3} y_{m,k}\beta_k^m, \tag{B.164}$$

where $y_{k,0} = -4\left(N+2\right)\phi_{kk}^3$,

$$y_{k,1} = 4\left(2N^3 + 7N^2 + 6N + 1\right)\phi_{kk}^2,$$

$$y_{k,2} = -\left(2N^5 + 24N^4 + 51N^3 + 45N^2 + 18N + 2\right)\phi_{kk},$$

$$y_{k,3} = \left(N^6 + 11N^5 + 22N^4 + 36N^3 + 34N^2 + 18N + 3\right). \tag{B.165}$$

We show (see ***Gumbel_N.nb***) that for any $N \geq 7$, $\phi_{kk} > 0$, $z_k^* < \frac{1}{5}\ln 2$ and

$$f\left(N\right)\phi_{kk} < \beta_k < \min\{f_{CS}(N,\phi_{kk}), \gamma(N,\phi_{kk},u_k^0)\},$$

the right-hand side of (B.164) is negative and $d_{CSk}\left(z_k^*, \beta_k, \phi_{kk}, N\right) > 0$ for all z_k^*, where $f_{CS}(N,\phi_{kk})$ is the largest real root of the third degree polynomial $\sum_{m=0}^{3} y_{m,k}\beta_k^m$. Thus, $\left.\frac{\partial CS_k^*}{\partial N}\right|_{\varphi_1=0} < 0$.

We next verify that the polynomial $\sum_{m=0}^{3} y_{m,k}\beta_k^m$ has three real roots and the largest real root, $f_{CS}(N,\phi_{kk})$, is linear in the externality ϕ_{kk} and can thus be expressed in separable form as $f_{CS}(N,\phi_{kk}) = f_{CS}(N)\phi_{kk}$. We define

$$\breve{t}_k := -\frac{(N+1)t\left(N\right)\phi_{kk}^2}{3\left(N^6 + 11N^5 + 22N^4 + 36N^3 + 34N^2 + 18N + 3\right)^2},$$
$$\breve{s}_k := -\frac{2s\left(N\right)\phi_{kk}^3}{27\left(N^6 + 11N^5 + 22N^4 + 36N^3 + 34N^2 + 18N + 3\right)^3}, \text{ and} \tag{B.166}$$
$$\breve{\Delta}_k := -\frac{4\delta\left(N\right)\phi_{kk}^6}{27\left(N^6 + 11N^5 + 22N^4 + 36N^3 + 34N^2 + 18N + 3\right)^4},$$

where $t(N) = \sum_{j=0}^{9} a_j N^j$, $s(N) = \sum_{j=0}^{15} b_j N^j$, $\delta(N) = \sum_{j=0}^{16} c_j N^j$ and

$$\{a_0, \ldots, a_9\} = \{-32, -328, -1124, -1516, -671, 577, 740, 364, 68, 4\},$$

$$\{b_0, \ldots, b_{15}\} = \{872, 10098, 45378, 102078, 118692, 39084, -73701, -98271, -22581, 50841, 61656, 34542, 11412, 2214, 216, 8\},$$

$$\{c_0, \ldots, c_{16}\} = \{-816, -9464, -42275, -92522, -97149, -4202, 113985, 130285, 40882, -40699, -51820\}$$
$$\cup \{-24692, -4259, 1118, 729, 136, 8\}.$$

From (B.166), it follows that for any $N \geq 7$, $\phi_{kk} > 0$, then $\breve{\Delta}_k < 0$. It thus follows that $\sum_{m=0}^{3} y_{m,k}\beta_k^m$ has three simple real roots given by

$$\breve{\alpha}_j := 2\sqrt{-\frac{\breve{t}_k}{3}} \cos\left(\frac{\theta + 2j\pi}{3}\right) - \frac{y_{2,k}}{3y_{3,k}}, \quad \text{for } j = 0, 1, 2,$$
$$0 < \theta < \pi, \text{ and } \cos(\theta) = \frac{-\breve{s}_k/2}{\sqrt{-(\breve{t}_k/3)^3}}. \tag{B.167}$$

From (B.165) and (B.167), for $\phi_{kk} < 0$, $\theta = \theta(N)$ is a function of N and is independent of ϕ_{kk}. Moreover, $\sqrt{-\frac{\breve{t}_k}{3}}$ and $\frac{y_{2,k}}{3y_{3,k}}$ are linear in ϕ_{kk}. Thus, for $\phi_{kk} < 0$, the three simple real roots of $\sum_{m=0}^{3} y_{m,k}\beta_k^m$ can be written as

$$\breve{\alpha}_j = w_j(N)\phi_{kk}. \tag{B.168}$$

We can thus write $f_{CS}(N, \phi_{kk}) = f_{CS}(N)\,\phi_{kk}$, where

$$f_{CS}(N) = \max_{j=0,1,2}\{w_j(N)\}. \tag{B.169}$$

Finally, we show that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, (i) and (ii) in Proposition 21 hold. From (B.131) and (B.132), $\partial(CS_k^*)/\partial N$ is a rational function w.r.t. (ϕ_{bs}, ϕ_{sb}). Moreover, at $(\phi_{bs}, \phi_{sb}) = (0, 0)$, $\partial(CS_k^*)/\partial N$ is given by (B.156). Thus, $\partial(CS_k^*)/\partial N$ is continuous w.r.t. (ϕ_{bs}, ϕ_{sb}) at $(0, 0)$. We conclude that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0) \subset \mathbb{R}^2$, (i) and (ii) in Proposition 21 hold true.

□

Proof of Proposition 22. We want to compute the following quantity when $\varphi_1 = 0$,

$$\frac{\partial \pi_k^*}{\partial N} = \frac{\partial p_k^*}{\partial N} x_k^* + p_k^* \frac{\partial x_k^*}{\partial N}. \tag{B.170}$$

Using equations (B.129), (B.131) and (B.149), we show (see *Gumbel_N.nb*) that

$$\left.\frac{\partial \pi_k^*}{\partial N}\right|_{\varphi_1=0} = \frac{n_{\pi k}\left(z_k^*, \beta_k, \phi_{kk}, N, u_k^0\right)}{d_{\pi k}\left(z_k^*, \beta_k, \phi_{kk}, N\right)}, \tag{B.171}$$

where $n_{\pi k}\left(z_k^*, \beta_k, \phi_{kk}, N, u_k^0\right)$ and $d_{\pi k}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ can be written as polynomials in $e^{z_k^*}$ in the following way,

$$\begin{aligned} n_{\pi k}\left(z_k^*, \beta_k, \phi_{kk}, N, u_k^0\right) &:= \sum_{m=2}^{7} n_{m,\pi k} e^{m z_k^*}, \qquad \text{and} \\ d_{\pi k}\left(z_k^*, \beta_k, \phi_{kk}, N\right) &:= \sum_{m=0}^{7} d_{m,\pi k} e^{m z_k^*}. \end{aligned} \tag{B.172}$$

Moreover, the coefficients $n_{m,\pi k}$ are as follows:

$$\begin{aligned}
n_{2,\pi k} &= \beta_k^3\left(u_k^0 + \beta_k z_k^*\right), \\
n_{3,\pi k} &= \beta_k^2\left(\beta_k^2\left((5N-2) z_k^* - 1\right) + \beta_k\left((5N-2) u_k^0 - 2 z_k^* \phi_{kk}\right) - 2 u_k^0 \phi_{kk}\right), \\
n_{4,\pi k} &= \beta_k\left[\beta_k^3\left(10N^2 z_k^* - 2N\left(4 z_k^* + 2\right) + z_k^*\right) + \beta_k^2\left(\left(10N^2 - 8N + 1\right) u_k^0 + \left(z_k^* + 1 - 6N z_k^*\right)\phi_{kk}\right)\right] \\
&+\beta_k\left[\beta_k\left(u_k^0\left(1-6N\right)\phi_{kk} + z_k^* \phi_{kk}^2\right) + u_k^0 \phi_{kk}^2\right], \\
n_{5,\pi k} &= \beta_k\left[\beta_k^2\left(N\left(10N^2 - 12N + 3\right) u_k^0 + \left(2N z_k^* + 4N - 1\right)\phi_{kk} - 6N^2 z_k^* \phi_{kk}\right)\right], \\
&+\beta_k\left[\beta_k^3 N\left(10N^2 z_k^* - 12N z_k^* - 6N + 3 z_k^*\right) + \beta_k \phi_{kk}\left(N\left(2-6N\right) u_k^0 + \left(N z_k^* + z_k^* + 2\right)\phi_{kk}\right) + \left(N+1\right) u_k^0 \phi_{kk}^2\right], \\
n_{6,\pi k} &= \beta_k\left[\beta_k^2\left(N^2\left(5N^2 - 8N + 3\right) u_k^0 + \left(N^2 z_k^* - 2N^3 z_k^* + 5N^2 - 2N\right)\phi_{kk}\right)\right], \\
&+\beta_k\left[\beta_k^3 N^2\left(5N^2 z_k^* - 2N\left(4 z_k^* + 2\right) + 3 z_k^*\right) + \beta_k \phi_{kk} N\left(-2N^2 u_k^0 + N u_k^0 + \left(z_k^* + 2\right)\phi_{kk}\right) + \left(N u_k^0 - 2\phi_{kk}\right)\phi_{kk}^2\right],
\end{aligned}$$

$$n_{7,\pi k} = \beta_k^3 N^2\left[\beta_k N\left(N^2 z_k^* - 2N z_k^* + z_k^* - N\right) + N u_k^0 + 2N\phi_{kk} - \phi_{kk} + N^3 u_k^0 - 2N^2 u_k^0\right]. \tag{B.173}$$

The coefficients $d_{m,\pi k}$ are given by

$$d_{0,\pi k} = \beta_k^3,$$

$$d_{1,\pi k} = \beta_k^2 \left(\beta_k \left(7N-1\right) - 4\phi_{kk}\right),$$

$$d_{2,\pi k} = \beta_k \left(\beta_k^2 \left(21N^2-7N+1\right) + 4\beta_k \left(1-5N\right)\phi_{kk} + 5\phi_{kk}^2\right),$$

$$d_{3,\pi k} = \beta_k^3 N \left(35N^2-20N+5\right) + \beta_k^2 \left(-40N^2+15N-1\right)\phi_{kk} + \beta_k \left(15N-3\right)\phi_{kk}^2 - 2\phi_{kk}^3,$$

$$d_{4,\pi k} = N \left(\beta_k^3 N \left(35N^2-30N+10\right) + \beta_k^2 \left(-40N^2+21N-3\right)\phi_{kk} + \beta_k \left(15N-5\right)\phi_{kk}^2 - 2\phi_{kk}^3\right),$$

$$d_{5,\pi k} = \beta_k N^2 \left(\beta_k^2 N \left(21N^2-25N+10\right) + \beta_k \left(-20N^2+13N-3\right)\phi_{kk} + \left(5N-1\right)\phi_{kk}^2\right),$$

$$d_{6,\pi k} = \beta_k N^3 \left(\beta_k^2 N \left(7N^2-11N+5\right) + \beta_k \left(-4N^2+3N-1\right)\phi_{kk} + \phi_{kk}^2\right),$$

$$d_{7,\pi k} = \beta_k^3 \left(N-1\right)^2 N^5. \tag{B.174}$$

Because the expressions determining $n_{\pi k}\left(z_k^*, \beta_k, \phi_{kk}, N, u_k^0\right)$ and $d_{\pi k}\left(z_k^*, \beta_k, \phi_{kk}, N\right)$ are complex, we focus on finding sufficient conditions for these expressions to have a specific sign for all z_k^*.

(i) Case $\left.\frac{\partial \pi_k^*}{\partial N}\right|_{\varphi_1=0} < 0$: We verify in the supplementary file *Gumbel_N.nb* that $n_{\pi k}$ and $d_{\pi k}$ (see (B.172)) are negative and positive, respectively, if either of the five conditions below, (a-i)-(a-v), hold:

(a-i) $\phi_{kk} < 0$, $N \geq 2$, $0 < \beta_k < f_\pi\left(N, \phi_{kk}\right)$ and $z_k^* < r_{\pi,z,1}\left(N, \phi_{kk}, u_k^0, \beta_k\right)$ where

$$r_{\pi,z,1}\left(N, \phi_{kk}, u_k^0, \beta_k\right) := \frac{\sum_{l=0}^{3} n_{\pi z,l}\beta_k^l}{\beta_k^2\left(N^2-2N^3\right)\phi_{kk} + \beta_k^3\left(5N^4-8N^3+3N^2\right) + \beta_k N \phi_{kk}^2} \tag{B.175}$$

with coefficients given by

$$\begin{gathered} n_{\pi z,0} = \phi_{kk}^2\left(-Nu_k^0 + 2\phi_{kk}\right), \\ n_{\pi z,1} = N\phi_{kk}\left(2N^2u_k^0 - Nu_k^0 - 2\phi_{kk}\right), \\ n_{\pi z,2} = N\left(-5N^3u_k^0 + 8N^2u_k^0 - 3Nu_k^0 - 5N\phi_{kk} + 2\phi_{kk}\right), \\ n_{\pi z,3} = 4N^3. \end{gathered} \tag{B.176}$$

Moreover, $f_\pi(N, \phi_{kk})$ is the largest real root of the third degree polynomial

$$4\beta_k^3 N^3 + \beta_k^2 (2 - 5N) N\phi_{kk} - 2\beta_k N\phi_{kk}^2 + 2\phi_{kk}^3 = 0. \tag{B.177}$$

We clarify below that the above polynomial has three real roots and that $f_\pi(N, \phi_{kk})$ is linear in ϕ_{kk} and can thus be expressed as $f_\pi(N, \phi_{kk}) = f_\pi(N)\phi_{kk}$.

(a-ii) $\phi_{kk} < 0$, $N \geq 2$, $\beta_k \geq f_\pi(N)\phi_{kk}$ and $z_k^* < -\frac{u_k^0}{\beta_k}$.

(a-iii) $\phi_{kk} = 0$, $N \geq 2$, $\beta_k > 0$ and $z_k^* < -\frac{u_k^0}{\beta_k}$.

(a-iv) $\phi_{kk} > 0$, $N \geq 2$, $f(N)\phi_{kk} < \beta_k \leq h_\pi(N)\phi_{kk}$ and $z_k^* < r_{\pi,z,2}(N, \phi_{kk}, u_k^0, \beta_k)$ where

$$h_\pi(N) := \frac{(2N-1)}{N^2}, \text{ and} \tag{B.178}$$

$$r_{\pi,z,2}\left(N, \phi_{kk}, u_k^0, \beta_k\right) := \frac{-N^3 u_k^0 + \beta_k N^2 + 2N^2 u_k^0 - N u_k^0 - 2N\phi_{kk} + \phi_{kk}}{\beta_k (N^3 - 2N^2 + N)}. \tag{B.179}$$

(a-v) $\phi_{kk} > 0$, $N \geq 2$, and $\beta_k > h_\pi(N)\phi_{kk}$ and $z_k^* < -\frac{u_k^0}{\beta_k}$.

For completeness, we quickly verify that as stated in (a-i), (B.177) has three real roots and that $f_\pi(N, \phi_{kk})$ is linear in ϕ_{kk}. We use Cardano's condition, for $\phi_{kk} < 0$, $\hat{\Delta}_k := (\hat{s}_k/2)^2 + (\hat{t}_k/3)^3 < 0$, where

$$\begin{aligned} \hat{t}_k &:= -\frac{(49N^2 - 20N + 4)\phi_{kk}^2}{48N^4}, \text{ and} \\ \hat{s}_k &:= \frac{(N+2)(127N^2 - 32N + 4)\phi_{kk}^3}{864N^6}. \end{aligned} \tag{B.180}$$

We note that for $\phi_{kk} < 0$, $f_\pi(N)$ satisfies

$$f_\pi(N) := \max_{j=0,1,2} \left\{ \frac{\hat{\alpha}_j}{\phi_{kk}} \right\}, \tag{B.181}$$

where $\hat{\alpha}_j$ is given by (B.139), after replacing (t_k, s_k) by $(\hat{t}_k, \hat{s}_k)$.

The five conditions (a-i)-(a-v) can be aggregated into the following single condition: $n_{\pi k}$ and $d_{\pi k}$ are negative and positive, respectively, if (β_k, ϕ_{kk}, N) satisfy (3.19) and $z_k^* <$

$g_{\pi,z}\left(N,\phi_{kk},u_k^0,\beta_k\right)$, where

$$g_{\pi,z}\left(N,\phi_{kk},u_k^0,\beta_k\right):=\begin{cases} r_{\pi,z,1}\left(N,\phi_{kk},u_k^0,\beta_k\right) & \text{if } \phi_{kk}<0 \text{ and } 0<\beta_k<f_\pi\left(N\right)\phi_{kk},\\ r_{\pi,z,2}\left(N,\phi_{kk},u_k^0,\beta_k\right) & \text{if } \phi_{kk}>0 \text{ and } f\left(N\right)\phi_{kk}<\beta_k\leq h_\pi\left(N\right)\phi_{kk},\\ -u_k^0/\beta_k & \text{if } \phi_{kk}=0 \text{ or } \left(\phi_{kk}<0 \text{ and } \beta_k\geq f_\pi\left(N\right)\phi_{kk}\right),\\ -u_k^0/\beta_k & \text{if } \left(\phi_{kk}>0 \text{ and } \beta_k>h_\pi\left(N\right)\phi_{kk}\right)\end{cases} \tag{B.182}$$

and the quantities $r_{\pi,z,1}\left(N,\phi_{kk},u_k^0,\beta_k\right)$, $f_\pi\left(N\right)$, $h_\pi\left(N\right)$ and $r_{\pi,z,2}\left(N,\phi_{kk},u_k^0,\beta_k\right)$ are given by (B.175), (B.181), (B.178) and (B.179), respectively.

(ii) Case $\frac{\partial\pi_k^*}{\partial N}|_{\varphi_1=0}>0$. We verify in the supplementary file *Gumbel_N.nb* that $n_{\pi k}$ and $d_{\pi k}$ (see (B.172)) are positive if either of the two conditions below, (b-i) or (b-ii), hold:

(b-i) $\phi_{kk}\leq 0$, $N\geq 2$, $\beta_k>0$ and $z_k^*>f_{\pi,z}\left(N,\phi_{kk},u_k^0,\beta_k\right)$, where

$$f_{\pi,z}\left(N,\phi_{kk},u_k^0,\beta_k\right):=r_{\pi,z,2}\left(N,\phi_{kk},u_k^0,\beta_k\right), \tag{B.183}$$

and $r_{\pi,z,2}\left(N,\phi_{kk},u_k^0,\beta_k\right)$ is given by (B.179).

(b-ii) $\phi_{kk}>0$, $N\geq 2$, $\beta_k>g_\pi\left(N,\phi_{kk}\right)$ and $z_k^*>f_{\pi,z}\left(N,\phi_{kk},u_k^0,\beta_k\right)$, where $g_\pi\left(N,\phi_{kk}\right)$ is the unique real root of the third degree polynomial

$$\beta_k^3N^3\left(N^2-1\right)+\beta_k^2N\left(-7N^3+10N^2-5N+1\right)\phi_{kk}+\beta_kN\left(6N^2-7N+3\right)\phi_{kk}^2-(2N^2-2N+1)\phi_{kk}^3. \tag{B.184}$$

We next verify that the above polynomial indeed has a unique real root and that $g_\pi\left(N,\phi_{kk}\right)$ is linear in ϕ_{kk} and can thus be expressed as $g_\pi\left(N,\phi_{kk}\right)=g_\pi\left(N\right)\phi_{kk}$. We use Cardano's condition, $\overline{\Delta}_k:=(\bar{s}_k/2)^2+(\bar{t}_k/3)^3>0$, where

$$\bar{t}_k:=-\frac{\left(31N^6-119N^5+179N^4-135N^3+54N^2-10N+1\right)\phi_{kk}^2}{3(N-1)^2N^4(N+1)^2}, \text{ and}$$

$$\bar{s}_k:=-\frac{\left(362N^9-2013N^8+4878N^7-6728N^6+5781N^5-3186N^4+1117N^3-237N^2+30N-2\right)\phi_{kk}^3}{27(N-1)^3N^6(N+1)^3}.$$

Next, we express this solution. For $\phi_{kk} > 0$, $g_\pi(N)$ satisfies

$$g_\pi(N) := \frac{\text{Car}(\overline{s}_k, \overline{\Delta}_k) - \frac{\overline{b}_2}{3}}{\phi_{kk}}, \tag{B.185}$$

where $\overline{b}_2 = -\frac{\left(7N^3 - 10N^2 + 5N - 1\right)\phi_{kk}}{(N-1)N^2(N+1)}$ and $\text{Car}(\cdot, \cdot)$ is given by (B.145).

Finally, we show that there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, (i) and (ii) in Proposition 22 hold. From (B.131) and (B.132), $\partial\pi_k^*/\partial N$ is a rational function w.r.t. (ϕ_{bs}, ϕ_{sb}). Moreover, at $(\phi_{bs}, \phi_{sb}) = (0, 0)$, the derivative $\partial\pi_k^*/\partial N$ is given by (B.171). Thus, $\partial\pi_k^*/\partial N$ is continuous w.r.t. (ϕ_{bs}, ϕ_{sb}) at $(0, 0)$. Therefore, there exists $\epsilon > 0$ such that for any $(\phi_{bs}, \phi_{sb}) \in B_\epsilon(0)$, cases (i) and (ii) in Proposition 22 hold true.

□

Appendix C

Additional Experiments Supporting Chapter 4

Section Appendix C.1 presents sensitivity analysis and exemplifies convergence paths. Section Appendix C.2 provides extra details on how to fit the univariate and bivariate functions of equation (4.19). Section Appendix C.3 provides additional numerical experiments on the dependence of the collusive level on β_k and $u_k^{(0)}$.

C.1 Sensitivity Analysis and Some Examples of Convergence Paths

We assume special parameterizations of the network externality matrix, Φ, and perform a sensitivity analysis for our simulation framework. For each specific Φ, we ran 500 simulations and classified the behavior in the last 5000 time steps of each run into one of five categories:

1. Symmetric 1-cycle (or 1-Sym), where platforms 1 and 2 repeatedly choose actions p_1 and p_2, respectively, and $p_1 = p_2$.

2. Asymmetric 1-cycle (or 1-Asym), where platforms 1 and 2 repeatedly choose actions p_1 and p_2, respectively, and $p_1 \neq p_2$.

3. Cycle of length 2-4 (or C2-4), where platforms 1 and 2 repeatedly choose actions that follow a pattern of length two to four. For example, the actions in a cycle of length 3 are $\{(p_1, p_2), (q_1, q_2), (r_1, r_2)\}$.

4. Cycle of length 5-8 (or C5-8), where the two platforms repeatedly choose actions that follow a pattern of length five to eight.

5. Cycle of length at least 9 (or C9), where the two platforms repeatedly choose actions that follow a pattern of length at least nine.

Let $k \in \{1, \ldots, 5\}$ index the five categories described above, and C_k denote the set of simulations within category k. For each $k \in \{1, \ldots, 5\}$ and a chosen Φ, we report the following measures:

- Frequency (Freq.) of the category among the 500 simulations.

- Average (Avg.) and standard deviation (S.D.) of the collusive level $\tilde{\Delta}$ within the category.

- Frequency of equilibrium, which quantifies the proportion of simulations in which the AI driven platforms learn the action that achieves maximal total future reward given that the opponent plays with the final Q-function. We compute this frequency, denoted by $Y_k \in [0, 1]$ for $k \in \{1, \ldots, 5\}$, according to the following three steps:

  1. Let Q_i^* be the final Q-function of platform $i \in \{1, 2\}$ in a given simulation $s \in C_k$. It is the numerical approximation to Q_i^* defined in (4.10), as explained in Section 4.3.2. We use this notation in defining other frequencies below. We estimate, without loss of generality, platform's 1 best response to Q_2^* using backward induction. For $\hat{T} = 10$, initiate

  $$Q_1^{(\hat{T})}(p, x) = \pi^{(1)}(p, \text{argmax}_{p'} Q_2^*(p', x)),$$

  where p is a vector of prices and x is the state variable.
  For $t \in \{1, \ldots, \hat{T}\}$, compute via backward induction

  $$Q_1^{(t-1)}(p, x) = \pi^{(1)}(p, \text{argmax}_{p'} Q_2^*(p', x)) + \delta \max_{p''} Q_1^{(t)}(p'', x'),$$

  where $x' = (p, \text{argmax}_{p'} Q_2^*(p', x))$. The matrix $Q_1^{(0)}$ is used as an approximation to platform's 1 best response to Q_2^* in the given simulation $s \in C_k$.

  2. Let P_s denote the set of states used in the convergence path of Q_i^* in the given simulation $s \in C_k$. We use this definition in defining the other frequencies

below. For each state $(p_1, p_2) \in P_s$, we form an indicator variable $X_s(p_1, p_2)$ that checks if the action that platform 1 takes using Q_1^* approximates the action that it would take using $Q_1^{(0)}$. Thus, for each $(p_1, p_2) \in P_s$, let $p^0 := \text{argmax}_{p'} Q_1^{(0)}(p', (p_1, p_2))$ and $p^* := \text{argmax}_{p'} Q_1^*(p', (p_1, p_2))$, then $X_s(p_1, p_2) = 1$ if and only if either $p^0 = p^*$ or p^0 and p^* are neighbors with distance at most 1 in the set of indices of the Q-matrix function.

3. Let $Y_k = \frac{1}{|C_k|} \sum_{s \in C_k} \frac{1}{|P_s|} \sum_{(p_1,p_2) \in P_s} X_s(p_1, p_2)$.

- Frequency of one-step equilibrium, which quantifies the proportion of simulations in which the AI driven platforms learn the action that achieves maximal one-step reward given that the opponent plays with the final Q-function. Note that the underlying learning involves no memory. We compute this frequency, denoted by $Y_k^{(\text{one})} \in [0, 1]$ for $k \in \{1, \ldots, 5\}$, according to the following three steps:

1. Let $Q_i^{(\text{one})}(p, x) = \pi^{(1)}(p, \text{argmax}_{p'} Q_2^*(p', x))$.
2. Let $p^{(\text{one})} := \text{argmax}_{p'} Q_1^{(\text{one})}(p', (p_1, p_2))$. We define $X_s^{(\text{one})}(p_1, p_2) = 1$ if and only if either $p^{(\text{one})} = p^*$ or $p^{(\text{one})}$ and p^* are neighbors with distance at most 1 in the set of indices of the Q-matrix function.
3. Let $Y_k^{(\text{one})} = \frac{1}{|C_k|} \sum_{s \in C_k} \frac{1}{|P_s|} \sum_{(p_1,p_2) \in P_s} X_s^{(\text{one})}(p_1, p_2)$.

- Frequency of converging back to the limiting cycle, which quantifies the proportion of simulations in which after one platform unilaterally deviates to the one-stage Nash equilibrium price, platforms return back to the limiting cycle. We compute this frequency, $Y_k^{(\text{b})} \in [0, 1]$ for $k \in \{1, \ldots, 5\}$, according to the following three steps:

1. At time $T+1$, set $p_{T+1}^{(1)} = (p_b^*, p_s^*)$, while $p_{T+1}^{(2)} \in \text{argmax}_p Q_2^*(p, (p_T^{(1)}, p_T^{(2)}))$. For $\tau \in \{2, \ldots, 101\}$ and $i \in \{1, 2\}$, let $p_{T+\tau}^{(i)} \in \text{argmax}_p Q_i^*(p, (p_{T+\tau-1}^{(1)}, p_{T+\tau-1}^{(2)}))$.
2. Let $p^{(\text{b}),(i)} := p_{T+101}^{(i)}$. We define $X_s^{(\text{b})}(p^{(\text{b}),(1)}, p^{(\text{b}),(2)}) = 1$ if and only if $(p^{(\text{b}),(1)}, p^{(\text{b}),(2)}) \in P_s$.
3. Let $Y_k^{(\text{b})} = \frac{1}{|C_k|} \sum_{s \in C_k} X_s^{(\text{b})}(p^{(\text{b}),(1)}, p^{(\text{b}),(2)})$.

- Frequency of converging back to the limiting cycle for both platforms, which quantifies the proportion of simulations in which after both platforms deviate to the one-stage Nash equilibrium price, they return back to the limiting cycle. This frequency is computed similarly to the one above, where the only difference is that $p^{(1)}_{T+1} = p^{(2)}_{T+1} = (p^*_b, p^*_s)$.

- Average and standard deviation of the Q-loss w.r.t. P_s. The Q-loss quantifies how close are the observed rewards on path P_s, $\sum_{\tau=0}^{100} \delta^\tau \pi^{(i)}_{t+\tau}$, to the optimal rewards given by the best response $Q^{(0)}_i$. We highlight that after convergence, the observed rewards should be pretty close in value to $\max_{p\in\mathcal{A}} Q^*_i(p,x)$ for each state $x \in P$. Let $T_0 = T - 5000$, we compute the Q-loss for a given simulation s as

$$\frac{1}{|P_s|} \sum_{t=T_0}^{T_0+|P_s|} \left| \sum_{\tau=0}^{100} \delta^\tau \pi^{(i)}_{t+\tau} - \max_{p\in\mathcal{A}} Q^0_i(p, x_t) \right|.$$

  The average and standard deviation are taken w.r.t. all simulations $s \in C_K$ for any $k \in \{1, \ldots, 5\}$.

- Average and standard deviation of the Q-loss w.r.t. all states. The Q-loss w.r.t. all states quantifies in average how close are the maximum values of $Q^*_i(\cdot, x)$ and $Q^0_i(\cdot, x)$, where the maximum is taken over all actions and the average over all states $x \in \mathcal{S}$. We compute the Q-loss w.r.t. all states for a given simulation s as

$$\frac{1}{|\mathcal{S}|} \sum_{x\in\mathcal{S}} \left| \max_{p\in\mathcal{A}} Q^*_i(p, x) - \max_{p\in\mathcal{A}} Q^0_i(p, x) \right|.$$

  The average and standard deviation are taken with respect to all simulations $s \in C_K$ for any $k \in \{1, \ldots, 5\}$. When the Q-loss w.r.t. all states equals 0, platforms' maximal rewards of the final Q-function yield the same maximal rewards of the best response Q-function. We interpret this as follows: the final Q-function exhibits behavior consistent with equilibrium off-path, also known as subgame perfection.

Tables C.1, C.2 and C.3 show the above measures for the following three respective choices of Φ: $\Phi_1 := [0, 0; 0, 0]$, $\Phi_2 := [1, 0; 0, 1]$ and $\Phi_3 := [0, 1; 1, 0]$. The other parameters are set as in the default setting, i.e., $\delta = 0.05$, $\beta_k = 1$, $u^{(0)}_k = -2$.

In all tables, the frequency of symmetric 1-cycles is relatively small, where its average is 5.8%. Even though this event could be considered rare, this category is important since within it we can find one-memory stationary equilibria, which are of great interest (see, e.g., Barlo et al. (2016) and Chica et al. (2024b)). Note that the frequency of asymmetric 1-

cycles is very small, where its average is 2.7%. We will thus omit interpretations of results for this category. The average of both frequencies of cycles of length 2-4 and 5-8 is 34% and the one for cycles of length at least 9 is 23.3%.

The average collusive level among all tables and categories, excluding 1-Asy, is approximately 29%. This value is only slightly below the baseline collusive level Δ_0.

For 1-Sym, the frequencies of equilibrium are 46%, 28% and 52% for Φ_1, Φ_2 and Φ_3, respectively. Thus, approximately half of the simulations within 1-Sym result in equilibrium behavior for Φ_1 and Φ_3, with a smaller value of 28% for Φ_2. These numbers suggest that in cases of zero externalities or non-zero cross-side externalities, equilibrium behavior is more likely than in cases of non-zero within-side externalities. Note that for the frequency of one-step equilibrium in 1-Sym, we still observe a larger proportion of one-step equilibrium for Φ_1 and Φ_3 compared to Φ_2.

For the categories C2-4, C5-8 and C9, the average frequencies of equilibrium are 28.8%, 42.9% and 49.4% for Φ_1, Φ_2 and Φ_3, respectively. Thus, on average 46% of the simulations with cycle lengths greater than one result in equilibrium behavior for Φ_2 and Φ_3, with a smaller value of 28.8% for Φ_1. These numerical results suggest that in cases of non-zero externalities and for cycles of lengths greater than one, equilibrium behavior is more likely than in the zero-externalities case. The above results for 1-Sym, C2-4, C5-8 and C9 suggest that in a sufficiently large percentage of cases, AI driven platforms exhibit behavior that is consistent with Nash equilibrium.

The frequencies of converging back to the limiting cycle for one platform in the 1-Sym category are 84%, 81% and 76% for Φ_1, Φ_2 and Φ_3, respectively. Thus, in at least 76% of the cases, platforms converge back to the limiting cycle after one unilateral price change to the Nash equilibrium price. For the categories C2-4, C5-8 and C9, the frequencies of converging back to the limiting cycle for one platform are 90%, 92% and 95% for Φ_1, Φ_2 and Φ_3, respectively. Note that the smallest frequency of converging back to the limiting cycle for one platform in the 1-Sym category is achieved at Φ_3, which suggests that non-zero cross-side externalities make it harder to sustain collusion. A similar result is shown by Ruhmer (2010) for a case of platform competition without AI agents. Nonetheless, this result does not hold for cycles of lengths greater than one, where the smallest frequency is achieved at Φ_1, i.e., at the zero-externalities case. As a general observation, the above results suggest that AI driven platforms can learn behavior that is consistent with tacit collusion. Note that similar behavior is found for the frequencies of converging back to the limiting cycle for both platforms.

The average Q-losses (on path) for all tables and categories, excluding 1-Asy, are all less than 0.019. This observation and the definition of the Q-loss imply that the observed rewards on path are sufficiently close to the optimal rewards given by the best response matrix. These results suggest that the likelihood of the final action achieving maximal total future reward is very high. Similarly, the average Q-losses w.r.t. all states for all categories, excluding 1-Asy, are less than 0.043. Thus, platforms' maximal rewards of the final Q-function are relatively close to the maximal rewards of the best response Q-function.[1] Furthermore, the Q-function exhibits behavior consistent with the equilibrium off-path or subgame perfection.

Metric	1-Sym	1-Asy	C2-4	C5-8	C9	All
Freq.	0.040	0.025	0.300	0.307	0.328	1.000
Avg. Collusive Level	0.273	0.326	0.294	0.301	0.295	0.297
S.D. of Collusive Level	0.082	0.069	0.063	0.049	0.042	0.054
Freq. of Eq.	0.462	0.0	0.279	0.296	0.29	0.288
Freq. of one-step Eq.	0.385	0.125	0.314	0.324	0.345	0.325
Freq. Conv. back (one)	0.846	0.750	0.907	0.909	0.896	0.898
Freq. Conv. back (both)	0.846	0.750	0.856	0.879	0.925	0.882
Avg. Q loss (on path)	0.003	0.006	0.005	0.005	0.005	0.005
S.D. of Q loss (on path)	0.001	0.004	0.003	0.002	0.001	0.002
Avg. Q loss (all)	0.009	0.009	0.009	0.009	0.009	0.009
S.D. of Q loss (all)	0.000	0.000	0.000	0.000	0.000	0.000

Table C.1: Sensitivity analysis with $\Phi = \Phi_1 := [0, 0; 0, 0]$.

[1]We note that achieving zero Q-loss is unrealistic for two reasons: the finite discretization of the action space and the unfeasible assumption that the optimal solution requires each platform to know the strategy of the other. Similarly, Calvano et al. (2020b) did not achieve zero Q-loss, despite assuming the simpler Bertrand model and using a more refined discretization.

Metric	1-Sym	1-Asy	C2-4	C5-8	C9	All
Freq.	0.084	0.013	0.355	0.368	0.179	1.000
Avg. Collusive Level	0.297	0.433	0.298	0.301	0.322	0.305
S.D. of Collusive Level	0.061	0.230	0.065	0.058	0.067	0.068
Freq. of Eq.	0.281	0.0	0.44	0.436	0.411	0.414
Freq. of one-step Eq.	0.063	0.000	0.366	0.326	0.320	0.313
Freq. Conv. back (one)	0.812	0.400	0.904	0.957	0.912	0.911
Freq. Conv. back (both)	0.812	0.400	0.889	0.943	0.897	0.897
Avg. Q loss (on path)	0.011	0.042	0.015	0.017	0.019	0.016
S.D. of Q loss (on path)	0.005	0.030	0.011	0.010	0.011	0.011
Avg. Q loss (all)	0.043	0.043	0.043	0.043	0.043	0.043
S.D. of Q loss (all)	0.001	0.002	0.001	0.001	0.001	0.001

Table C.2: Sensitivity analysis with $\Phi = \Phi_2 := [1, 0; 0, 1]$.

Metric	1-Sym	1-Asy	C2-4	C5-8	C9	All
Freq.	0.049	0.042	0.364	0.353	0.193	1.000
Avg. Collusive Level	0.264	0.282	0.279	0.265	0.287	0.275
S.D. of Collusive Level	0.057	0.079	0.075	0.053	0.061	0.065
Freq. of Eq.	0.524	0.336	0.533	0.499	0.451	0.497
Freq. of one-step Eq.	0.245	0.342	0.369	0.365	0.321	0.351
Freq. Conv. back (one)	0.762	0.500	0.930	0.967	0.976	0.926
Freq. Conv. back (both)	0.667	0.333	0.904	0.954	0.964	0.898
Avg. Q loss (on path)	0.009	0.019	0.013	0.012	0.015	0.013
S.D. of Q loss (on path)	0.004	0.010	0.009	0.006	0.009	0.008
Avg. Q loss (all)	0.038	0.038	0.038	0.038	0.038	0.038
S.D. of Q loss (all)	0.001	0.001	0.001	0.001	0.001	0.001

Table C.3: Sensitivity analysis with $\Phi = \Phi_3 := [0, 1; 1, 0]$.

Figure C.1 shows four examples of convergence paths, one from C2-4 (top left panel), one from C5-8 (top right panel) and two from C9 (bottom left and right panels), where $\Phi = [0, 1; 1, 0]$. The horizontal blue dashed lines at the bottom and top portion of each panel represent the Nash and Collusion equilibrium prices, respectively, i.e., the lines at $p_b^* = p_s^*$ and $p_b^C = p_s^C$, respectively. Note that buyer and seller prices are equal due to the symmetric choice of Φ. The vertical blue dotted lines represent where a cycle ends, while the blue and orange curves represent the buyer and seller prices, respectively, chosen by platform 1 on the equilibrium path. The top left panel shows a convergence path of length four, in which the buyer's price starts above the Nash equilibrium price. Then, in two steps, it reaches the Nash equilibrium price, followed by a sudden increase to the starting price. This price pattern is similar to an Edgeworth cycle, where prices start well above the Nash equilibrium price, then slowly converge to the Nash equilibrium, followed by a sudden increase to the initial high price (see, .e.g., Maskin and Tirole (1988)). Note that the seller's price oscillates between two levels above the Nash equilibrium price. The top right and bottom left and right subfigures show more intricate patterns. However, a common feature among them is that prices oscillate between the Nash equilibrium price and a level higher than the Nash price. This indicates that platforms learn behavior consistent with punishment and reward strategies.

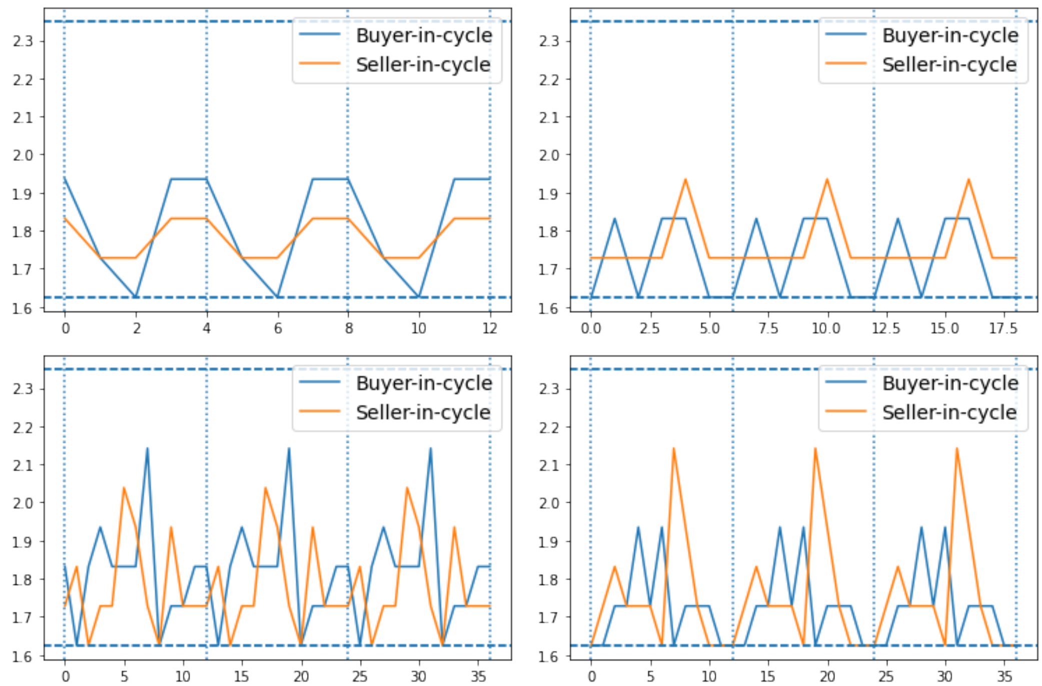


Figure C.1: Examples of convergence paths from categories C2-4 (top left panel), C5-8 (top right panel) and C9 (bottom left and right panels), where $\Phi = [0, 1; 1, 0]$. The horizontal blue dashed lines at the bottom and top of each panel represent the lines at $p_b^* = p_s^*$ and $p_b^C = p_s^C$, respectively. The vertical blue dotted lines indicate where a cycle ends. The blue and orange curves represent the buyer and seller prices, respectively, chosen by platform 1.

Finally, Figure C.2 illustrates four scenarios depicting the convergence path of platform 1 after both platforms revert to the limiting cycle following a deviation to the one-stage Nash equilibrium price by both platforms. The top left and right panels present an example for Φ_1, while the bottom left and right panels provide examples for Φ_2 and Φ_3, respectively. In all four cases, the horizontal blue dashed lines at the bottom and top of each panel represent the lines at $p_b^* = p_s^*$ and $p_b^C = p_s^C$, respectively. The vertical blue dotted lines indicate where each cycle starts and ends. The blue and orange curves represent the buyer and seller prices, respectively, chosen by platform 1 immediately after the deviation to the Nash equilibrium price. The green and red curves represent platform 1's buyer and seller prices once they revert to the limiting cycle. In all four panels, we observe that in less than ten steps, platform 1's prices revert to the limiting cycle (the leftmost vertical dotted line appears before step 10). The top left panel shows that after one deviation to the Nash equilibrium price by both platforms, platform 1's prices slowly increase above the Nash equilibrium price, followed by a small decrease and then a convergence to a price above the Nash equilibrium price. Similar behavior is observed in the top right panel. The bottom left and right panels show more intricate patterns. However, as mentioned earlier, in both figures, prices converge back to the limiting cycle after 9 steps. The patterns they follow after reaching the limiting cycle indicate that firms oscillate between the Nash equilibrium price and a price higher than the Nash equilibrium price.

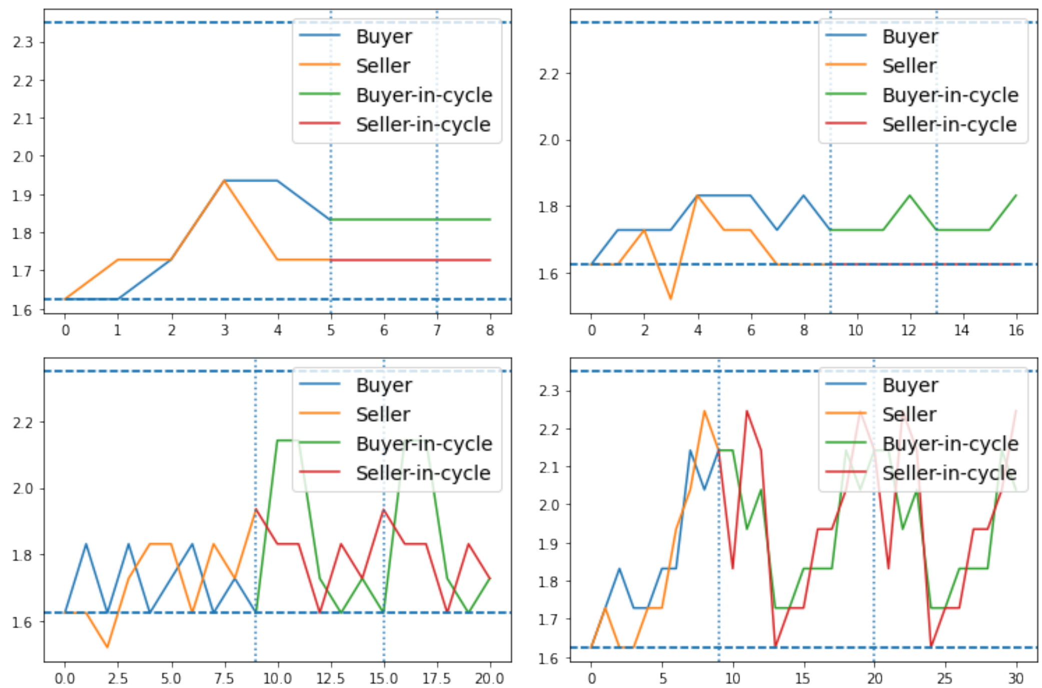


Figure C.2: Four examples of convergence paths after both platforms revert to the limiting cycle following a deviation to the one-stage Nash equilibrium price by both platforms. There are two examples for Φ_1 (top left and right panels), one for Φ_2 (bottom left), and one for Φ_3 (bottom right). The blue and orange curves represent the buyer and seller prices, respectively, chosen by platform 1 immediately after the deviation to the Nash equilibrium price. The green and red curves represent platform 1's buyer and seller prices once they revert to the limiting cycle.

C.2 Fitting a Sequence of Non-linear Functions

We provide the details of how to fit the univariate and bivariate functions in (4.19) in order to describe how the collusive level depends on the externality matrix Φ via an additive model. We first let Δ_0 equal the sample mean of $\tilde{\Delta}$. In order to find the four univariate functions $f_{bb}(\phi_{bb}), f_{ss}(\phi_{ss}), f_{bs}(\phi_{bs})$ and $f_{sb}(\phi_{sb})$, we fit them sequentially using residuals. We consider all the twenty-four permutations of 4 elements, represented by bb, ss, bs, sb. Without loss of generality, we describe the fitting procedure for a given permutation $o := (bb, ss, bs, sb)$, where the same procedure is followed for all other permutations. We follow a regression setting with the predictor ϕ_{bb} and the response

$$y_{bb} := \tilde{\Delta} - \Delta_0,$$

and we apply XGBoost, which is a nonlinear regression method, to fit $\hat{f}_{bb}^{(o)}(\phi_{bb})$ that approximates y_{bb}. For the above o, we proceed from bb to ss as follows. We define

$$y_{ss} := y_{bb} - \hat{f}_{bb}^{(o)}(\phi_{bb})$$

and consider a regression setting with y_{ss} as response and ϕ_{ss} as predictor. We apply XGBoost to approximate y_{ss} by $\hat{f}_{ss}^{(o)}(\phi_{ss})$. Similarly, we transition from ss to bs using the residuals

$$y_{bs} := y_{ss} - \hat{f}_{ss}^{(o)}(\phi_{ss})$$

and fitting $\hat{f}_{bs}^{(o)}(\phi_{bs})$ to approximate y_{bs} by XGBoost, and we transition from bs to sb using the residual

$$y_{sb} := y_{bs} - \hat{f}_{bs}^{(o)}(\phi_{ss})$$

and fitting $\hat{f}_{sb}^{(o)}(\phi_{sb})$ to approximate y_{sb} by XGBoost. We repeat this process for the rest of the twenty-three permutations and thus obtain twenty-four approximation for the functions $f_{bb}(\phi_{bb})$, $f_{ss}(\phi_{ss})$, $f_{bs}(\phi_{bs})$ and $f_{sb}(\phi_{sb})$, specified in (4.19). We average over the twenty-four approximations to obtain the estimator $\hat{f}_{bb}(\phi_{bb})$, $\hat{f}_{ss}(\phi_{ss})$, $\hat{f}_{bs}(\phi_{bs})$ and $\hat{f}_{sb}(\phi_{sb})$.

Next, we apply the procedure described above for the six bivariate functions in (4.19), such as $f_{bb,ss}(\phi_{bb}, \phi_{ss})$ and $f_{bs,sb}(\phi_{bs}, \phi_{sb})$. For a given permutation of the six pairs of variables, e.g.,

$$o = ((bb, ss), (bs, sb), (bb, sb), (bb, bs), (ss, sb), (ss, bs)),$$

we follow the same procedure introduced above to iteratively fit the bivariate functions using XGBoost. In the first iteration, the response variable is

$$y_1 := \tilde{\Delta} - \Big(\Delta_0 + \hat{f}_{bb}(\phi_{bb}) + \hat{f}_{ss}(\phi_{ss}) + \hat{f}_{bs}(\phi_{bs}) + \hat{f}_{sb}(\phi_{sb})\Big)$$

and for the above permutation o there are two predictors ϕ_{bb} and ϕ_{ss}. XGBoost then approximates y_1 by the function $\hat{f}^{(o)}(\phi_{bb}, \phi_{ss})$. In the following iterations, the response is the corresponding residual and the predictors are the two corresponding variables. For example, assuming the permutation o, in the second iteration the response variable is

$$y_1 - \hat{f}^{(o)}(\phi_{bb}, \phi_{ss})$$

and the predictors are ϕ_{bs} and ϕ_{sb}. We average the obtained approximations over all 720 possible permutations of these six bivariate functions to approximate each bivariate function.

C.3 Other Simulation Results

We provide Figures C.3 and C.4, which complement Figures 4.8 and 4.9, respectively, with additional choices of the externality matrix.

A common trend in both Figures 4.8 and C.3 is that the collusive level decreases as β_k increases.

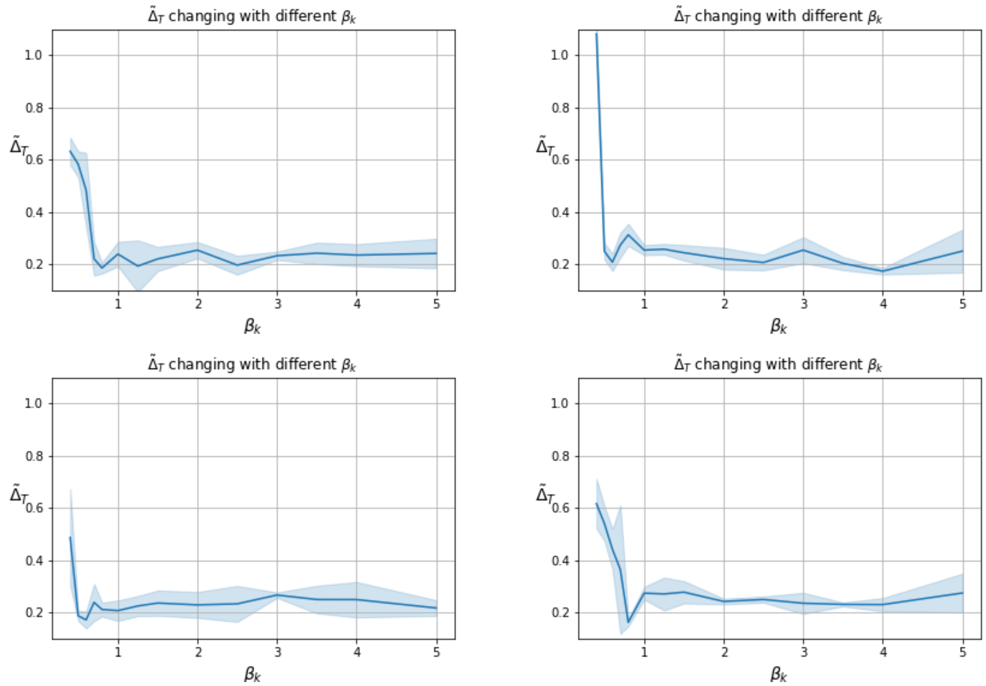


Figure C.3: The collusive level with varying $\beta_k \in [0.2, 5]$; Top left: $\Phi = [0, 1; 1, 0]$; Top right: $\Phi = [0, -1; -1, 0]$; Bottom left: $\Phi = [0, 1; 0, 0]$; Bottom right: $\Phi = [1, 0; 0, -1]$.

Section 4.4.3 summarizes the main common observation to both Figure 4.9 and Figure C.4.

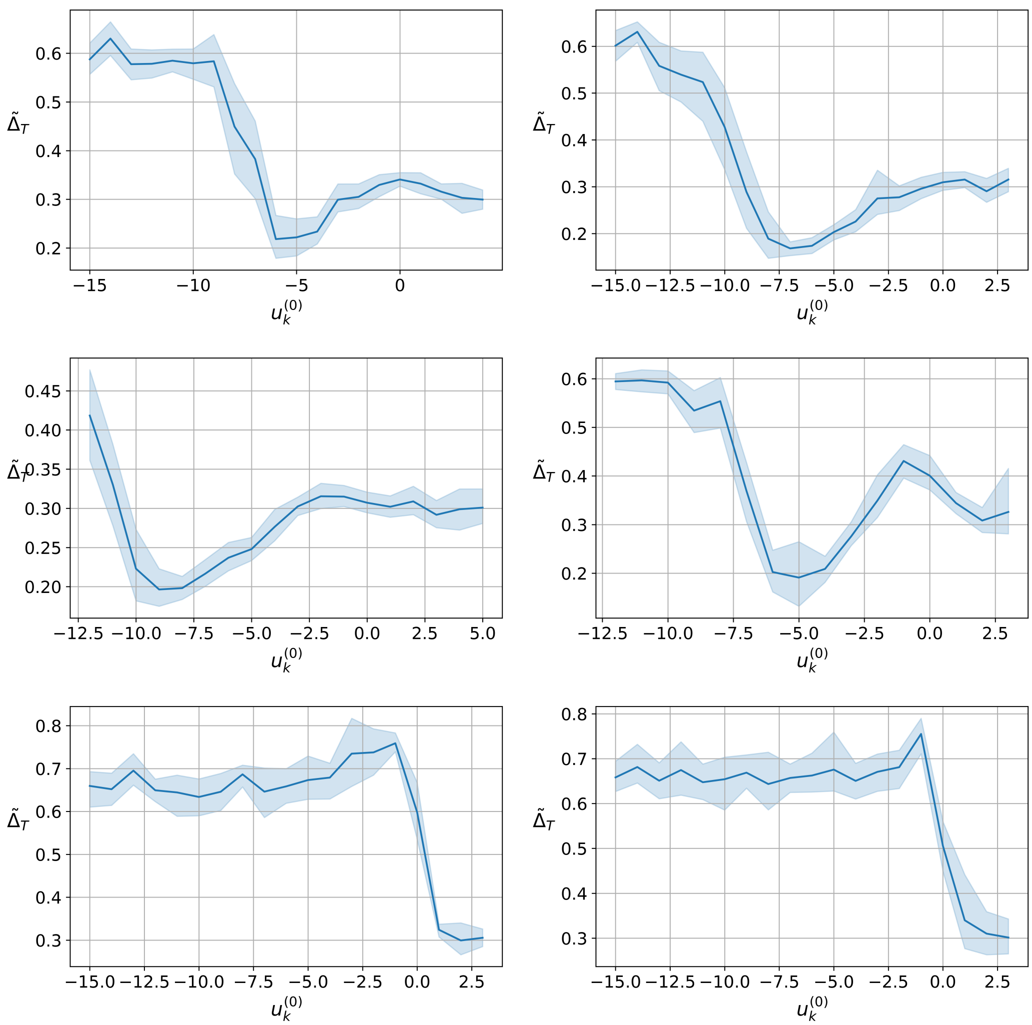


Figure C.4: The collusive level with varying $u_k^{(0)}$, for $k \in \{b, s\}$, where $\Phi = [1, 0; 0, 0]$ (top left), $\Phi = [0, 1; 0, 0]$ (top right), $\Phi = [0, 1; -1, 0]$ (middle left), $\Phi = [1, 0; 0, -1]$ (middle right), $\Phi = [2, 0; 0, 2]$ (bottom left), and $\Phi = [0, 2; 2, 0]$ (bottom right).