Dynamic Oligopoly Games with Private Markovian Dynamics

Yi Ouyang, Hamidreza Tavafoghi and Demosthenis Teneketzis

***Abstract*— We analyze a dynamic oligopoly model with strategic sellers and buyers/consumers over a finite horizon. Each seller has private information described by a finite-state Markov process; the Markov processes describing the sellers' information are mutually independent. At the beginning of each time/stage t the sellers simultaneously post the prices for their good; subsequently, consumers make their buying decisions; finally, after the buyers' decisions are made, a public signal, indicating the buyers' consumption experience from each seller's good becomes available and the whole process moves to stage $t+1$. The sellers' prices, the buyers' decisions and the signal indicating the buyers' consumption experience are common knowledge among buyers and sellers. This dynamic oligopoly model arises in online shopping and dynamic spectrum sharing markets. The model gives rise to a stochastic dynamic game with asymmetric information. Using ideas from the common information approach (developed in [1] for decentralized decision-making), we prove the existence of common information based equilibria. We obtain a sequential decomposition of the game and we provide a backward induction algorithm to determine common information-based equilibria that are perfect Bayesian equilibria. We illustrate our results with an example.**

I. Introduction

A. Background and Motivation

Stochastic dynamic games arise in many socio-technological systems that consist of many strategic decision makers (agents). In dynamic games with symmetric information all the agents share the same information and each agent makes decisions anticipating other agents' strategies. This class of dynamic games has been extensively studied in the literature (see [2–5] and references therein). An appropriate solution concept for this class of games is sub-game perfect equilibrium (SPE) which consists of a strategy profile of agents that must satisfy *sequential rationality* [2]. The common history in dynamic games with symmetric information can be utilized to provide a sequential decomposition of the dynamic game. The common history (or a function of it) serves as an information state and SPE can be computed through backward induction.

In dynamic games with asymmetric information agents have different observations over time, and consequently different local history. As a result, each agent needs to anticipate the other agents' strategy and to form a belief about the other agents' local information. Therefore, perfect Bayesian equilibrium (PBE) is an appropriate solution concept for this class of games. PBE consists of a pair of strategy profiles and beliefs for all agents that jointly must satisfy *sequential rationality* and *consistency* [2]. In games with asymmetric information a decomposition similar to that of games with symmetric information is not possible in general. This is because the computation of any belief at time t depends, in general, on the strategy of all agents up to time $t-1$. As a result, only special instances of stochastic dynamic games with asymmetric information have been studied in the literature (see [6–12] and references therein).

In this paper, we consider a dynamic oligopoly model with asymmetric information and Markovian dynamics. We define a class of PBE and provide a sequential decomposition of the game through an appropriate choice of information state using ideas from the common information approach developed in [1] for decentralized decision-making. The proposed equilibrium and the associated decomposition resemble Markov perfect equilibrium (MPE) defined in [13] for dynamic games with symmetric information.

Dynamic oligopoly games have been studied in economics. Bergemann et al. [14] study a dynamic oligopoly game with symmetric information where the quality of goods is unknown and is revealed over time through public consumption experiences of the buyers; the solution concept is MPE. A repeated oligopoly game with asymmetric information is investigated in [15], [16] with private production costs that are described by iid processes; the solution concept is public perfect equilibrium [12]. In this paper, we consider a dynamic oligopoly model where each seller has private information about the quality of his product; this information is described by a Markov process. Such a dynamic oligopoly model results in a non-zero sum stochastic dynamic game with asymmetric information.

The dynamic oligopoly model investigated in this paper arises in several socio-technological areas. Such an area is the growing online shopping market. In such a market, the buyers' shopping decisions are based on the products' posted prices and the posted consumer reviews about their consumption experiences [17], [18]. A product's quality is the seller's private information that varies over time due to changes in technology and other unpredictable factors. It is widely known that the public consumer reviews, known as *word of mouth* [19], have great impact on the the consumers' buying decisions and, consequently, on the sellers' pricing strategies.

Another area is dynamic spectrum markets. It has been argued that the static long-term allocation of the spectrum results in inefficient and underutilization of spectrum [20]. A dynamic market architecture, regulated by FCC, has been

This research was supported in part by NSF grants CCF-1111061 and CNS-1238962.

Y. Ouyang, H. Tavafoghi and D. Teneketzis are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI (e-mail: ouyangyi@umich.edu; tavaf@umich.edu; teneket@umich.edu).

proposed as a solution to this inefficiency. In this market architecture primary users (service providers) can rent/lease spectrum (possibly along with access points) to secondary users for limited time and at particular locations [21], [22]. In such markets the primary users have private information about the quality and vacancy of their own channels. This asymmetry of of information has not been addressed in the literature so far. Existing work (see [23–26] and references in) has primarily investigated dynamic spectrum markets with symmetric information.

The problems investigated in [6–10] are the most closely related to our problem, but they are different from our problem. The authors of [8–10] analyze zero-sum Markovian models with asymmetric information. Our paper as well as [6], [7] use a common information-based methodology to establish the existence of common information-based equilibria, and achieve a sequential decomposition of the dynamic game that leads to a backward induction algorithm that determines such equilibria. The key difference between our problem and those in [6], [7] is in the information structure. The information structure in [6], [7] is such that the agents' common information based belief about the system state is policy independent. Therefore, in [6], [7] there is no signaling effect [2], [4] and one only needs to determine a strategy profile that satisfies the sequential rationality condition. The information structure in our problem results in a common information-based belief that is policy dependent. As a result, there is signaling involved and we need to determine both strategy profiles and belief systems that jointly satisfy sequential rationality and consistency conditions.

B. Contribution

The key contributions of the paper are: (i) The existence of common information-based equilibria for a class of stochastic dynamic games with asymmetric information where the common information-based beliefs are policy dependent; (2) The sequential decomposition of stochastic dynamic oligopoly games with the asymmetric information structure described in section II. This decomposition provides a backward induction algorithm to find common information-based equilibria that are PBE. The results are illustrated by an example.

C. Organization

The paper is organized as follows. We introduce the model and formulate our problem in section II. In Section III, we analyze the problem and present our results. We provide an example that illustrates our results in section IV. We conclude in section V. Due to space limitation, the proofs of all the technical results of the paper have been omitted. They can be found in the Appendix.

D. Notation

Random variables are denoted by upper case letters, their realization by the corresponding lower case letter. In general, subscripts are used as time index while superscripts are used to index sellers and buyers. For time indices $t_1 \leq t_2$, $X_{t_1:t_2}$ (resp. $f_{t_1:t_2}(\cdot)$) is the short hand notation for the variables $(X_{t_1}, X_{t_1+1}, ..., X_{t_2})$ (resp. functions $(f_{t_1}(\cdot), \ldots, f_{t_1}(\cdot))$). When we consider the variables (resp. functions) for all time, we drop the subscript and use X to denote $X_{1:T}$ (resp. $f(\cdot)$ to denote $f_{1:T}(\cdot)$). For variables $X_t^1, \ldots, X_t^N$ (resp. functions $f_t^1(\cdot), \ldots, f_t^N(\cdot)$), we use $X_t := (X_t^1, \ldots, X_t^N)$ (resp. $f_t(\cdot) := (f_t^1(\cdot), \ldots, f_t^N(\cdot))$) to denote the vector of the set of variables (resp. functions), and $X_t^{-n} := (X_t^1, \ldots, X_t^{n-1}, X_t^{n+1}, \ldots, X_t^N)$ (resp. $f_t^{-n}(\cdot) := (f_t^1(\cdot), \ldots, f_t^{n-1}(\cdot), f_t^{n+1}(\cdot), \ldots, f_t^N(\cdot))$) to denote all the variables (resp. functions) except that of the agent indexed by n. $\mathbb{P}(\cdot)$ and $\mathbb{E}(\cdot)$ are the probability and the expectation, respectively, of an event. For a set $\mathcal{X}$, $\Delta(\mathcal{X})$ denotes the set of all beliefs/distributions on $\mathcal{X}$. For random variables X, Y with realizations x, y, $\mathbb{P}(x|y) := \mathbb{P}(X = x|Y = y)$ and $\mathbb{E}(X|y) := \mathbb{E}(X|Y = y)$. For a strategy g and a belief (probability mass function) π, we use $\mathbb{P}_\pi^g(\cdot)$ (resp. $\mathbb{E}_\pi^g(\cdot)$) to indicate that the probability (resp. expectation) depends on the choice of g and π.

II. System Model and Problem Formulation

Consider a dynamic oligopoly game between N sellers indexed by $n \in \mathcal{N} = \{1, 2, \cdots, N\}$ and M buyers indexed by $m \in \mathcal{M} = \{1, 2, \cdots, M\}$ over a time period $\mathcal{T} = \{1, 2, \cdots, T\}$. At time $t \in \mathcal{T}$ the state of the system is given by $(X_t^1, X_t^2, \cdots, X_t^N)$, where $X_t^n \in \mathcal{X}$ denotes the state of seller n which is privately observed by seller n. The private state X_t^n has Markovian dynamics with matrix of transition probabilities given by $Q^n = \{Q^n(i, j), i, j \in \mathcal{X}\}$ for $n \in \mathcal{N}$.

At the beginning of time t, each seller $n, n \in \mathcal{N}$, selects a price $P_t^n \in \mathcal{P}$ at which he sells a unit of his good. After the prices $P_t := (P_t^1, P_t^2, \ldots, P_t^N)$ are announced, each buyer chooses the amount of good that he wants to buy from each seller at time t. Let $D_t^m \in \mathcal{D}^N, m \in \mathcal{M}$, denote buyer m's decision at time t. Then, $D_t^m := (D_t^m(1), \ldots, D_t^m(N))$, and $D_t^m(n) = k$ means that buyer m buys k units of good from seller n. Let $D_t := (D_t^1, \ldots, D_t^M)$ denote the buyers' decisions at time t. The decisions D_t are observed by both the buyers and the sellers. After the buyers' decisions are made, each buyer receives the procured goods at time t. The consumption experience of good n at time t, denoted by $Y_t^n \in \mathcal{Y}$, is given by

$$\mathbb{P}(y_t^n|x_t^n) = Q^y(x_t^n, y_t^n) \text{ for all } y_t^n \in \mathcal{Y}, x_t^n \in \mathcal{X}. \quad (1)$$

Let $Y_t := (Y_t^1, Y_t^2, \cdots, Y_t^N)$; Y_t is observed by the sellers and the buyers at time t. The time-ordering of events from time t to $t + 1$ is shown in Fig. 1. Double arrows (blue) indicate sellers' and buyers' decisions; thin arrows indicate the realizations of the states of the Markov processes and the consumption experiences; and dashed arrows indicate observations of sellers and buyers. Note that t_+ is used to denote the time when the buyers make decisions.

For each $n \in \mathcal{N}$, seller n's instantaneous revenue at time t is given by

$$\phi_t^{n,S}(P_t, D_t) := (P_t^n - c) \sum_{m \in \mathcal{N}} D_t^m(n) \quad (2)$$

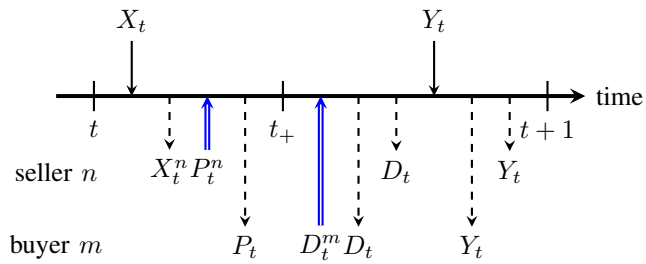

Fig. 1. The time-ordering of events.

where c denote the fixed marginal production cost. For each $m \in \mathcal{M}$, buyer m's instantaneous utility at time t is given by

$$\phi_t^{m,B}(Y_t, D_t, P_t) := v(Y_t, D_t) - \sum_{n \in \mathcal{N}} D_t^m(n) P_t^n, \quad (3)$$

where $v(Y_t, D_t)$ denotes his utility from using the procured goods with consumption experience Y_t which depends on the collective decisions D_t of the buyers. Note that the dependency of the buyer' utility on the aggregate decision D_t can capture complementarity among the goods and externality among the buyers.

We make the following assumptions

Assumption 1. *The dynamic oligopoly game is a finite game. i.e. $\mathcal{X}$, $\mathcal{P}$, $\mathcal{D}$ and $\mathcal{Y}$ are all finite sets.*

Assumption 2. *The Markov chains describing the evolution of the state processes, and the consumption experiences Y_t conditional on the state X_t of the sellers are all mutually independent.*

Let $H_t^{n,S}$ denote the history of observations of seller n at time t, and $H_t^{m,B}$ denote the history of observations of buyer m at time t. The histories $H_t^{n,S}$ and $H_t^{m,B}$ are given by (consult Fig 1)

$$H_t^{n,S} = \{X_{1:t}^n, P_{1:t-1}, D_{1:t-1}, Y_{1:t-1}\} \text{ for } n \in \mathcal{N}, \quad (4)$$

$$H_t^{m,B} = H_{t+}^c := \{P_{1:t}, D_{1:t-1}, Y_{1:t-1}\} \text{ for } m \in \mathcal{M}. \quad (5)$$

Let $\mathcal{H}_t^{n,S}$ and $\mathcal{H}_t^{m,B}$ denote the space of possible histories of seller n and buyer m, respectively, at time t. Let $\mathcal{H}_t := \cup_{n \in \mathcal{N}, m \in \mathcal{M}} \mathcal{H}_t^{n,S} \cup \mathcal{H}_t^{m,B}$ and $\mathcal{H} := \cup_{t \in \mathcal{T}} \mathcal{H}_t$ denote the set of all possible histories of the buyers and the sellers over time horizon T. A behavioral strategy for seller n at time t is a mapping $g_t^{n,S} : \mathcal{H}_t^{n,S} \mapsto \Delta(\mathcal{P})$ such that

$$\mathbb{P}(P_t^n = p_t^n) = g_t^{n,S}(h_t^{n,S})(p_t^n) \text{ for all } p_t^n \in \mathcal{P}. \quad (6)$$

A behavioral strategy for buyer m at time t is a mapping $g_t^{m,B} : \mathcal{H}_t^{m,B} \mapsto \Delta(\mathcal{D})$ such that

$$\mathbb{P}(D_t^m = d_t^m) = g_t^{m,B}(h_t^{m,B})(d_t^m) \text{ for all } d_t^m \in \mathcal{D}. \quad (7)$$

We define $g^{n,S} := g_{1:T}^{n,S}$ and $g^{m,B} := g_{1:T}^{m,B}$, and call $g := (g^{1,S}, \ldots, g^{N,S}, g^{1,B}, \ldots, g^{M,B})$ a strategy profile. Let $\mathcal{G}_t^{n,S}$ and $\mathcal{G}_t^{m,B}$ denote the set of all possible strategies of seller n and buyer m, respectively, at time t.

Let $\mu := (\mu_1, \mu_2 \ldots, \mu_T)$ be a belief system where $\mu_t : \mathcal{H}_t \mapsto \Delta(\mathcal{X}^{t \times N})$ for any $t \in \mathcal{T}$; for every history $h_t^{n,S} \in \mathcal{H}_t$ (resp. $h_t^{m,B} \in \mathcal{H}_t$) $\mu_t(h_t^{n,S})$ (resp. $\mu_t(h_t^{m,B})$) defines a distribution/belief on $X_{1:t}$ for seller n (resp. for buyer m).

For each $n \in \mathcal{N}$, seller n's objective is to maximize his expected (under the belief system μ) profit described by

$$\mathbb{E}_\mu^g \left[\sum_{t=1}^T \phi_t^{n,S}(P_t, D_t) \right] \quad (8)$$

where $\phi_t^{n,S}(P_t, D_t)$ is given (2). For each $m \in \mathcal{M}$, buyer m's objective is to maximize his expected (under the belief system μ) utility described by

$$\mathbb{E}_\mu^g \left[\sum_{t=1}^T \phi_t^{m,B}(Y_t, D_t, P_t) \right] \quad (9)$$

where $\phi_t^{m,B}(Y_t, D_t, P_t)$ is given by (3).

A pair (g, μ) satisfies *sequential rationality* if for every $n \in \mathcal{N}$, $g_{t:T}^{n,S}$ is a solution to

$$\sup_{g_{t:T}'^{n,S} \in \mathcal{G}_{t:T}^{n,S}} \mathbb{E}_\mu^{g_{t:T}'^{n,S}, g^{-n,S}, g^B} \left[\sum_{\tau=t}^T \phi_t^{n,S}(P_t, D_t) | h_t^{n,S} \right] \quad (10)$$

for every $t \in \mathcal{T}$ and every history $h_t^{n,S} \in \mathcal{H}_t^{n,S}$, and for every $m \in \mathcal{M}$, $g_{t:T}^{m,B}$ is a solution to

$$\sup_{g_{t:T}'^{m,B} \in \mathcal{G}_{t:T}^{m,B}} \mathbb{E}_\mu^{g^S, g_{t:T}'^{m,B}, g^{-m,B}} \left[\sum_{\tau=t}^T \phi_t^{m,B}(Y_t, D_t, P_t) | h_t^{m,B} \right], \quad (11)$$

for every $t \in \mathcal{T}$ and for every history $h_t^{m,B} \in \mathcal{H}_t^{m,B}$.

A pair (g, μ) satisfies *consistency* if for every $t \in \mathcal{T}$, $\mu_t(h_{t+1}^{n,S})$ (resp. $\mu_t(h_{t+1}^{m,B})$) is determined by Bayes' rule

$$\mu_t(h_{t+1}^{n,S})(x_{1:t+1}) = \frac{\mathbb{P}^g_{\mu_t(h_t^{n,S})}(x_{1:t+1}, p_t, d_t, y_t | h_t^{n,S})}{\sum_{x'_{1:t+1} \in \mathcal{X}^{t+1} : x'^n_{t+1} = x^n_{t+1}} \mathbb{P}^g_{\mu_t(h_t^{n,S})}(x'_{1:t+1}, p_t, d_t, y_t | h_t^{n,S})} \quad (12)$$

for every history $h_{t+1}^{n,S} = (h_t^{n,S}, p_t, d_t, y_t, x_{t+1}^n) \in \mathcal{H}_{t+1}^{n,S}$ of seller n, $n \in \mathcal{N}$ (respectively,

$$\mu_t(h_{t+1}^{m,B})(x_{1:t+1}) = \frac{\mathbb{P}^g_{\mu_t(h_t^{m,B})}(x_{1:t+1}, p_{t+1}, d_t, y_t | h_t^{m,B})}{\sum_{x'_{1:t+1} \in \mathcal{X}^{t+1}} \mathbb{P}^g_{\mu_t(h_t^{m,B})}(x'_{1:t+1}, p_{t+1}, d_t, y_t | h_t^{m,B})} \quad (13)$$

for every history $h_{t+1}^{m,B} = (h_t^{m,B}, p_{t+1}, d_t, y_t) \in \mathcal{H}_{t+1}^{m,B}$ of buyer m, $m \in \mathcal{M}$).

A pair (g, μ) is called a *perfect Bayesian equilibrium* (PBE) if it satisfies *sequential rationality* and *consistency*.[1]

We wish to determine PBE of the dynamic game with asymmetric information defined by the dynamic oligopoly model described in this section.

[1]Note that the dynamic oligopoly model defined in this paper satisfies the conditions of [27] [2, ch. 8], because the Markov processes describing the private information of the sellers are mutually independent. Therefore, the set of PBEs and the set of sequential equilibria of the game defined in this paper are the same.

III. Analysis

The main results of this paper are described by Theorem 1 and 2 presented in Section III-B. The proof of Theorem 1 and 2, along with the proofs of all the intermediate results required to verify the theorems' assertions can be found in the Appendix. To establish Theorems 1 and 2 we proceed as follows.

A. Preliminaries

The common history H_t^c at the beginning of time t is

$$H_t^c = \{P_{1:t-1}, D_{1:t-1}, Y_{1:t-1}\}. \tag{14}$$

The common history $H_{t_+}^c$ at time t_+, the time when buyers make decisions, is

$$H_{t_+}^c = \{P_{1:t}, D_{1:t-1}, Y_{1:t-1}\}. \tag{15}$$

Let $\mathcal{H}^c$ denote the space of all possible common histories. Note that $\mathcal{H}_t^{m,B} = \mathcal{H}_{t_+}^c$ and $h_t^{m,B} = h_{t_+}^c$ for all $m \in \mathcal{M}$.

Based on common histories, we define common-information based (CIB) belief systems. A map $\gamma : \mathcal{H}^c \mapsto \cup_{t\in\mathcal{T}}\Delta(\mathcal{X}^N)$ is called a CIB belief system. A set of functions $\psi = \{\psi_t^n, \psi_{t_+}^n, n \in \mathcal{N}, t \in \mathcal{T}\}$, where $\psi_t^n : \Delta(\mathcal{X}^N) \times \mathcal{P} \mapsto \Delta(\mathcal{X})$ and $\psi_{t_+}^n : \Delta(\mathcal{X}^N)\times\mathcal{Y} \mapsto \Delta(\mathcal{X})$, is called a CIB update rule. From any CIB update rule ψ, we can construct a CIB belief system $\gamma_\psi : \mathcal{H}^c \mapsto \cup_{t\in\mathcal{T}}\Delta(\mathcal{X}^N)$ by the following inductive construction:

1) $\gamma_\psi(h_1^c)(x_1) = \mathbb{P}(x_1) = \prod_{n\in\mathcal{N}} \mathbb{P}(x_1^n) \quad \forall x_1 \in \mathcal{X}^N$.
2) At time t_+, after $\gamma_\psi(h_t^c)$ is defined, set

$$\gamma_\psi(h_{t_+}^c)(x_t^n) := \psi_t^n(\gamma_\psi(h_t^c), p_t^n)(x_t^n), \tag{16}$$

$$\gamma_\psi(h_{t_+}^c)(x_t) := \prod_{n=1}^N \gamma_\psi(h_{t_+}^c)(x_t^n), \tag{17}$$

for all $x_t \in \mathcal{X}^N$.

3) At time $t+1$, after $\gamma_\psi(h_{t+}^c)$ is defined, set

$$\gamma_\psi(h_{t+1}^c)(x_{t+1}^n) := \psi_{t_+}^n(\gamma_\psi(h_{t_+}^c), y_t^n)(x_{t+1}^n), \tag{18}$$

$$\gamma_\psi(h_{t+1}^c)(x_{t+1}) := \prod_{n=1}^N \gamma_\psi(h_{t+1}^c)(x_{t+1}^n), \tag{19}$$

for all $x_{t+1} \in \mathcal{X}^N$.

For the CIB belief system γ_ψ, we use $\Pi_t^{\gamma_\psi}$ and $\Pi_{t_+}^{\gamma_\psi}$ to denote the belief, under γ_ψ, on X_t conditional on the common histories at time t and time t_+, respectively; that is,

$$\Pi_t^{\gamma_\psi} := \gamma_\psi(H_t^c) \in \Delta(\mathcal{X}^N), \tag{20}$$

$$\Pi_{t_+}^{\gamma_\psi} := \gamma_\psi(H_{t_+}^c) \in \Delta(\mathcal{X}^N). \tag{21}$$

We also define the marginal beliefs on X_t^n at time t and time t_+, respectively, as

$$\Pi_t^{n,\gamma_\psi}(x_t^n) := \gamma_\psi(H_t^c)(x_t^n) \quad \forall x_t^n \in \mathcal{X}, \tag{22}$$

$$\Pi_{t_+}^{n,\gamma_\psi}(x_t^n) := \gamma_\psi(H_{t_+}^c)(x_t^n) \quad \forall x_t^n \in \mathcal{X}. \tag{23}$$

Given a CIB belief system γ_ψ, we consider CIB strategies where each seller n makes his decision at time t based on X_t^n and $\Pi_t^{\gamma_\psi}$, and each buyer m makes his decision at time t_+ based on P_t and $\Pi_{t_+}^{\gamma_\psi}$. We call a set of functions $\lambda = \{\lambda_t^{n,S}, \lambda_t^{m,B}, n \in \mathcal{N}, m \in \mathcal{M}, t \in \mathcal{T}\}$, where $\lambda_t^{n,S} : \mathcal{X} \times \Delta(\mathcal{X}^N) \mapsto \Delta(\mathcal{P})$ and $\lambda_t^{m,B} : \mathcal{P} \times \Delta(\mathcal{X}^N) \mapsto \Delta(\mathcal{D})$, a CIB strategy profile.

In analogy with PBE, we define consistent beliefs for CIB strategies.

Definition 1 (consistency)**.** *For a given CIB strategy profile λ, we call a CIB update rule ψ consistent with λ if, for all $n \in \mathcal{N}$,*

$$\begin{aligned}&\psi_t^n(\pi_t, p_t^n)(x_t^n)\\ =&\frac{\lambda_t^{n,S}(x_t^n, \pi_t)(p_t^n)\pi_t^n(x_t^n)}{\sum_{x_t'^n\in\mathcal{X}} \lambda_t^{n,S}(x_t'^n, \pi_t)(p_t^n)\pi_t^n(x_t'^n)}\end{aligned} \tag{24}$$

when the denominator of (24) *is non-zero*;

$$\begin{aligned}&\psi_{t_+}^n(\pi_{t_+}, y_t^n)(x_{t+1}^n)\\ =&\frac{\sum_{x_t^n\in\mathcal{X}} Q^n(x_t^n, x_{t+1}^n)Q^y(x_t^n, y_t^n)\pi_{t_+}^n(x_t^n)}{\sum_{x_t'^n\in\mathcal{X}} Q^y(x_t'^n, y_t^n)\pi_{t_+}^n(x_t'^n)}\end{aligned} \tag{25}$$

when the denominator of (25) *is non-zero*.

Note that when the denominator of (24) and (25) are zero, $\psi_t^n(\cdot)$ and $\psi_{t_+}^n(\cdot)$ can be arbitrarily defined and consistency is still satisfied.

The following lemma connects a CIB strategy profile along with its consistent CIB belief to a strategy profile along with its consistent belief for the dynamic oligopoly game.

Lemma 1. *If λ is a CIB strategy profile and ψ is a CIB update rule consistent with λ, there exists a pair, denoted by $(g, \mu) = f(\lambda, \psi)$, such that $g \in \mathcal{G}$ is a strategy profile with*

$$g_t^{n,S}(h_t^{n,S}) := \lambda_t^{n,S}(x_t^n, \gamma_\psi(h_t^c)) = \lambda_t^{n,S}(x_t^n, \pi_t^{\gamma_\psi}) \tag{26}$$

for all $n \in \mathcal{N}$,

$$g_t^{m,B}(h_{t_+}^c) := \lambda_t^{m,B}(p_t, \gamma_\psi(h_{t_+}^c)) = \lambda_t^{m,B}(p_t, \pi_{t_+}^{\gamma_\psi}), \tag{27}$$

for all $m \in \mathcal{M}$, and $\mu : \mathcal{H} \mapsto \cup_{t\in\mathcal{T}}\Delta(\mathcal{X}^{t\times N})$, is a belief system such that, for all $h_t^{n,S}, h_{t_+}^c \in \mathcal{H}$ and for all $x_t \in \mathcal{X}^N$,

$$\mu(h_t^{n,S})(x_t) = \mathbf{1}_{\{x_t^n|h_t^{n,S}\}} \prod_{k\neq n} \gamma_\psi(h_t^c)(x_t^k), \tag{28}$$

$$\mu(h_{t_+}^c)(x_t) = \prod_{k\in\mathcal{N}} \gamma_\psi(h_{t_+}^c)(x_t^k), \tag{29}$$

where $\mathbf{1}_{\{x_t^n|h_t^{n,S}\}}$ denotes the indicator that $X_t^n = x_t^n$ in the history $h_t^{n,S}$. Furthermore, μ is consistent with g.

Note that, (28) implies that for any seller n, his private state X_t^n is independent of X_t^{-n} under μ for any history $h_t^{n,S} \in \mathcal{H}_t^{n,S}$.

B. Common-Information Based Equilibria

We focus on common-information based equilibria defined below.

Definition 2. *A pair (λ^*, ψ^*) of a CIB strategy profile λ^* and a CIB update rule ψ^* is called a common-information based equilibrium (CIB equilibrium) if ψ^* is consistent with*

λ^* and the pair $(g^*, \mu^*) = f(\lambda^*, \psi^*)$ defined in Lemma 1 forms a PBE.

The following lemma plays a crucial role in establishing the main results of this paper.

Lemma 2 (Closeness of CIB strategies)**.** *Suppose λ is a CIB strategy profile and ψ is a CIB update rule consistent with λ. If every buyer m uses the strategy $\lambda^{m,B}$ along with the belief generated by ψ, and every seller $k \neq n$ uses the strategy $\lambda^{k,S}$ along with the belief generated by ψ, then, there exists a CIB strategy $\tilde{\lambda}^{n,S}$ that is a best response for seller n under the belief generated by ψ. Furthermore, if every seller n uses the strategy $\lambda^{n,S}$ along with the belief generated by ψ, and every buyer $k \neq m$ uses the strategy $\lambda^{k,B}$ along with the belief generated by ψ, then, there exists a CIB strategy $\tilde{\lambda}^{m,B}$ that is a best response for buyer m under the belief generated by ψ.*

Lemma 2 says that the set of CIB strategies is closed under the best response mapping. Lemma 2 allows us to restrict attention to the set of CIB strategies and to search for CIB equilibria. We show that the best response mapping has a fixed point within the set of CIB strategies and beliefs; thus, we establish the existence of CIB equilibria.

Theorem 1. *The dynamic oligopoly game defined in Section II has at least one CIB equilibrium.*

In order to compute CIB equilibria, for a pair of CIB strategy profile λ and CIB update rule ψ (not necessary consistent with λ), we define, recursively, a set of functions

$$V(\lambda, \psi) := \left\{ V_t^{n,S}(\cdot), V_t^{m,B}(\cdot), V_{t_+}^{n,S}(\cdot), V_{t_+}^{m,B}(\cdot), t \in \mathcal{T} \right\} \quad (30)$$

as follows.
$V_{T+1}^{n,S}(\cdot) := 0 \quad \forall n \in \mathcal{N}$, $V_{T+1}^{m,B}(\cdot) := 0 \quad \forall m \in \mathcal{M}$.
For each $t_+, t \in \mathcal{T}$, for all $n \in \mathcal{N}$ and for all $m \in \mathcal{M}$,

$$V_{t_+}^{n,S}(\pi_{t_+}, p_t, x_t^n) := \mathbb{E}_{\pi_{t_+}}^{\lambda_t^B} \Big[\phi_t^{n,S}(p_t, D_t) + V_{t+1}^{n,S}(\psi_{t_+}(\pi_{t_+}, Y_t), X_{t+1}^n) | \pi_{t_+}, p_t, x_t^n \Big] \quad (31)$$

$$V_{t_+}^{m,B}(\pi_{t_+}, p_t) := \mathbb{E}_{\pi_{t_+}}^{\lambda_t^B} \Big[\phi_t^{m,B}(Y_t, p_t, D_t) + V_{t+1}^{m,B}(\psi_{t_+}(\pi_{t_+}, Y_t)) | \pi_{t_+}, p_t \Big] \quad (32)$$

where π_{t_+} denotes the CIB distribution of X_t at t_+ in the above expectation.
For each $t, t \in \mathcal{T}$, for all $n \in \mathcal{N}$ and for all $m \in \mathcal{M}$,

$$V_t^{n,S}(\pi_t, x_t^n) := \mathbb{E}_{\pi_t}^{\lambda_t^S} \left[V_{t_+}^{n,S}(\psi_t(\pi_t, P_t), P_t, x_t^n) | \pi_t, x_t^n \right] \quad (33)$$

$$V_t^{m,B}(\pi_t) := \mathbb{E}_{\pi_t}^{\lambda_t^S} \left[V_{t_+}^{m,B}(\psi_t(\pi_t, P_t), P_t) | \pi_t \right]. \quad (34)$$

where π_t denotes the CIB distribution of X_t at t in the above expectation.

Using the set of functions $V(\lambda, \psi)$, we provide a method/algorithm to sequentially compute CIB equilibria in the following theorem.

Theorem 2. *A pair (λ^*, ψ^*) of a CIB strategy profile λ^* and a CIB update rule ψ^* is a CIB equilibrium if and only if ψ^* is consistent with λ^* and λ^* solves the dynamic program for functions $V^* = V(\lambda^*, \psi^*)$. That is,*

1) *For all $n \in \mathcal{N}$, for any $x_t^n \in \mathcal{X}$ and $\pi_t \in \Delta(\mathcal{X}^N)$, any $p_t^{*n} \in \mathcal{P}$, such that $\lambda_t^{*n,S}(x_t^n, \pi_t)(p_t^{*n}) > 0$, should satisfy*

$$p_t^{*n} \in \arg\max_{p_t^n \in \mathcal{P}} \Big\{ \mathbb{E}_{\pi_t}^{\lambda_t^{*-n,S}} \Big[V_{t_+}^{*n,S}(\psi_t^*(\pi_t, (p_t^n, P_t^{-n})), (p_t^n, P_t^{-n}), x_t^n) | \pi_t, x_t^n \Big] \Big\}. \quad (35)$$

2) *For all $m \in \mathcal{M}$, for any $p_t \in \mathcal{P}^N$ and $\pi_{t_+} \in \Delta(\mathcal{X}^N)$, any $d_t^{*m} \in \mathcal{D}$, such that $\lambda_t^{*m,B}(p_t, \pi_{t_+})(d_t^{*m}) > 0$, should satisfy*

$$d_t^{*m} \in \arg\max_{d_t^m \in \mathcal{D}} \Big\{ \mathbb{E}_{\pi_{t_+}}^{\lambda_t^{*-m,B}} \Big[\phi_t^{m,B}(Y_t, p_t, (d_t^m, D_t^{-m})) + V_{t+1}^{*m,B}(\psi_{t_+}^*(\pi_{t_+}, Y_t)) | p_t, \pi_{t_+} \Big] \Big\}. \quad (36)$$

Any CIB equilibrium (λ^*, ψ^*) determined by algorithm of Theorem 2 gives rise to a PEB $(g^*, \mu^*) = f(\lambda^*, \psi^*)$ via the result of Lemma 1.

We illustrate our results with the following example.

IV. An Example

Consider an oligopoly game with $N = 2$ sellers and $M = 10$ buyers over a time horizon $T = 2$. The state of each seller's service is either high or low, *i.e.* $\mathcal{X} = \{H, L\}$, with transition probabilities given by

$$Q^n(H, H) = q_H, Q^n(H, L) = 1 - q_H, \quad (37)$$

$$Q^n(L, H) = q_L, Q^n(H, L) = 1 - q_L \quad (38)$$

for both the buyers; the initial distribution on X_1^1 and X_1^2 are given by π_1^1 and π_1^2, respectively. The quality of the good Y_t^n is assumed to be either $y > 0$ or 0. The conditional distribution of Y_t^n is given by

$$\mathbf{P}(Y_t^n = y) = q_H^y \text{ when } X_t^n = H, \quad (39)$$

$$\mathbf{P}(Y_t^n = y) = q_L^y \text{ when } X_t^n = L, \quad (40)$$

where $q_H^y > q_L^y > 0$. Consequently

$$\mathbb{E}[Y_t^n | X_t^n = H] = q_H^y y, \quad \mathbb{E}[Y_t^n | X_t^n = L] = q_L^y y. \quad (41)$$

Let $\mathcal{P} = \{h, l\}$, that is, each seller can set either a high price h or a low price l. We assume that each buyer buys at most one unit of the good from either of the sellers, i.e. $\mathcal{D} = \{(0,0), e^1, e^2\}$ where $e^1 := (1,0), e^2 := (0,1)$.

The instantaneous utility of buyer m is given by

$$\phi_t^{m,B}(Y_t, P_t, D_t) := \sum_{n=1,2} D_t^m(n) \left(Y_t^n - P_t^n - \sum_{k \neq m} D_t^k(n) \right) \quad (42)$$

where the term $\sum_{k \neq m} D_t^k(n)$ captures the negative consumption externality created by other buyers (*e.g.* delay in delivery in online shopping; interference in spectrum

markets). The instantaneous revenue of seller n are given by

$$\phi_t^{n,S}(P_t, D_t) = P_t^n \sum_{m\in\mathcal{M}} D_t^m(n). \tag{43}$$

Since the state is either H or L, any common belief on X_t^n at time t (resp. t_+) can be described by a scalar π_t^n (resp. $\pi_{t_+}^n$) that denotes $\mathbb{P}(X_t^n = H|h_t^c)$ (resp. $\mathbb{P}(X_t^n = H|h_{t_+}^c)$).

Using Theorem 2, we compute a CIB equilibrium of the dynamic oligopoly game described above.

A. Computation of a CIB Equilibrium

In the computation, we use the following parameters: $q_H = 0.8, q_L = 0.3, q_H^y = 0.6, q_L^y = 0.1, y = 10, h = 10, l = 8$. Due to space limit, see the Appendix for more details in this section.

First consider time $t_+, t = 1, 2$. Suppose that the agents use a CIB strategy profile λ and a CIB update rule ψ. Then the dynamic problem (36) for buyer m at each time $t_+, t = 1, 2$, becomes

$$\begin{aligned}
&\arg\max_{d_t^m\in\mathcal{D}} \Big\{\mathbb{E}_{\pi_{t_+}}^{\lambda_t^{-m,B}} \Big[\phi_t^{m,B}(Y_t, p_t, (d_t^m, D_t^{-m})) \\
&\quad + V_{t+1}^{m,B}(\psi_{t_+}(\pi_{t_+}, Y_t))|p_t, \pi_{t_+}\Big]\Big\} \\
&= \arg\max_{d_t^m\in\mathcal{D}} \Big\{\mathbb{E}_{\pi_{t_+}}^{\lambda_t^{-m,B}} \Big[\phi_t^{m,B}(Y_t, p_t, (d_t^m, D_t^{-m}))|p_t, \pi_{t_+}\Big]\Big\} \\
&\quad + \mathbb{E}_{\pi_{t_+}} \Big[V_{t+1}^{m,B}(\psi_{t_+}(\pi_{t_+}, Y_t))|p_t, \pi_{t_+}\Big].
\end{aligned} \tag{44}$$

Note that the term $\mathbb{E}_{\pi_{t_+}} \Big[V_{t+1}^{m,B}(\psi_{t_+}(\pi_{t_+}, Y_t))|p_t, \pi_{t_+}\Big]$ does not depend on D_t. Therefore, a CIB strategy $\lambda_t^B(\cdot)$ solves (44) if for any $p_t \in \mathcal{P}$ and $\pi_{t_+} \in \Delta(\mathcal{X}^N)$, $\{\lambda_t^{m,B}(p_t, \pi_{t_+})(d_t^m), m \in \mathcal{M}\}$ is a Bayesian Nash equilibrium (BNE) of the game where each buyer $m, m \in \mathcal{M}$ has the objective:

$$\begin{aligned}
&\max_{d_t^m\in\mathcal{D}} \Big\{\mathbb{E}_{\pi_{t_+}} \Big[\phi_t^{m,B}(Y_t, p_t, (d_t^m, d_t^k))|p_t, \pi_{t_+}\Big]\Big\} = \\
&\max_{d_t^m\in\mathcal{D}} \left\{\sum_{n=1,2} d_t^m(n) \left(\pi_{t_+}^n y_D + q_L^y y - p_t^n - \sum_{k\neq m} d_t^k(n)\right)\right\}
\end{aligned} \tag{45}$$

where $y_D := q_H^y y - q_L^y y$. It can be shown that the above game has a symmetric equilibrium given by

$$\begin{aligned}
&\mathbb{P}(D_t^m(n) = 1) \\
&= \frac{1}{2(M-1)} \left(y_D(\pi_{t_+}^n - \pi_{t_+}^{-n}) - p_t^n + p_t^{-n}\right) + \frac{1}{2}
\end{aligned} \tag{46}$$

for $t = 1, 2$ and $n = 1, 2$. Therefore, we define the behavioral strategy $\lambda_t^{*m,B}(p_t, \pi_{t_+})$ of buyer $m, m \in \mathcal{M}$, at time $t_+, t = 1, 2$, by

$$\begin{aligned}
&\lambda_t^{*m,B}(p_t, \pi_{t_+})(e^n) \\
&= \frac{1}{2(M-1)} \left(y_D(\pi_{t_+}^n - \pi_{t_+}^{-n}) - p_t^n + p_t^{-n}\right) + \frac{1}{2}.
\end{aligned} \tag{47}$$

Now suppose the buyers use strategy λ^{*B} and the sellers use a (behavioral) CIB strategy λ^S given by

$$\lambda_1^{n,S}(x_t^n, \pi_1)(h) = \beta_1^n(\pi_1) \text{ for both } x_t^n = H \text{ or } L, \tag{48}$$

$$\lambda_2^{n,S}(H, \pi_2)(h) = \beta_2^n(H, \pi_2), \tag{49}$$

$$\lambda_2^{n,S}(L, \pi_2)(h) = \beta_2^n(L, \pi_2), \tag{50}$$

where $\beta_1 = \{\beta_1^n(\pi_t^1) \in [0, 1], n = 1, 2\}$ and $\beta_2 = \{\beta_2^n(H, \pi_2), \beta_2^n(L, \pi_2) \in [0, 1], n = 1, 2\}$ are functions that describe the probabilities in the behavioral strategies $\lambda_n^{n,S}, n = 1, 2$. The CIB strategy λ^S can be interpreted as follows. At time $t = 1$ the sellers make pricing decisions based only on the CIB belief π_1. At time $t = 2$ the sellers make pricing decisions based on their private information and the the CIB belief π_2.

For any β_2 consider a CIB update rule ψ^{β_2} given by

$$\psi_1^{n,\beta_2}(\pi_1, p_1^n) = \pi_1^n \text{ for both } p_1^n = h \text{ or } l, \tag{51}$$

$$\begin{aligned}
&\psi_{1_+}^{n,\beta_2}(\pi_{1_+}, y) \\
&= \frac{\pi_{1_+}^n q_H^y q_H + (1 - \pi_{1_+}^n) q_L^y q_L}{\pi_{1_+}^n q_H^y + (1 - \pi_{1_+}^n) q_L^y},
\end{aligned} \tag{52}$$

$$\begin{aligned}
&\psi_{1_+}^{n,\beta_2}(\pi_{1_+}, 0) \\
&= \frac{\pi_{1_+}^n (1 - q_H^y) q_H + (1 - \pi_{1_+}^n)(1 - q_L^y) q_L}{\pi_{1_+}^n (1 - q_H^y) + (1 - \pi_{1_+}^n)(1 - q_L^y)},
\end{aligned} \tag{53}$$

for $n = 1, 2$ at time $t = 1$; at time 2 when $\beta_2^n(H, \pi_2) = \beta_2^n(L, \pi_2)$

$$\psi_2^{n,\beta_2}(\pi_2, p_2^n) = \pi_2^n \text{ for both } p_2^n = h \text{ or } l; \tag{54}$$

when $\beta_2^n(H, \pi_2) \neq \beta_2^n(L, \pi_2)$

$$\psi_2^{n,\beta_2}(\pi_2, h) = \frac{\pi_2^n \beta_2^n(H, \pi_2)}{\pi_2^n \beta_2^n(H, \pi_2) + (1 - \pi_2^n)\beta_2^n(L, \pi_2)}, \tag{55}$$

$$\begin{aligned}
&\psi_2^{n,\beta_2}(\pi_2, l) = \\
&\frac{\pi_2^n (1 - \beta_2^n(H, \pi_2))}{\pi_2^n (1 - \beta_2^n(H, \pi_2)) + (1 - \pi_t^n)(1 - \beta_2^n(L, \pi_2))}
\end{aligned} \tag{56}$$

for $n = 1, 2$. The CIB update rule ψ^{β_2} is consistent with the CIB strategy profile $(\lambda^S, \lambda^{*B})$.

Then, under the CIB strategy profile $(\lambda^S, \lambda^{*B})$ and the CIB update rule ψ^{β_2}, the expected instantaneous revenue of seller $n, n = 1, 2$, at time 2 becomes

$$p_2^n \mathbb{E}_{\pi_2}^{(\lambda^S, \lambda^{*B})} \left[\sum_{m\in\mathcal{M}} D_2^m(n)\right] = r_2^{n,\beta_2}(\pi_2, p_2^n) \tag{57}$$

where

$$\begin{aligned}
r_2^{n,\beta_2}(\pi_2, p_2^n) := &\frac{M}{2(M-1)} p_2^n [y_D(\psi_2^{n,\beta_2}(\pi_2, p_2^n) - \pi_2^{-n}) \\
&- p_2^n + \bar{p}_2^{-n}(\pi_2, \beta_2) + (M-1)],
\end{aligned} \tag{58}$$

and

$$\begin{aligned}
&\bar{p}_2^n(\pi_2, \beta_2) := (\pi_2^n \beta_2^n(H, \pi_2) + (1 - \pi_2^n)\beta_2^n(L, \pi_2))h \\
&+ (\pi_2^n (1 - \beta_2^n(H, \pi_2)) + (1 - \pi_2^n)(1 - \beta_2^n(L, \pi_2)))l
\end{aligned} \tag{59}$$

for $n = 1, 2$. There exists a BNE $\lambda_2^{*S}(\cdot)$ at time $t = 2$ for the game where each seller has the objective given by (57),

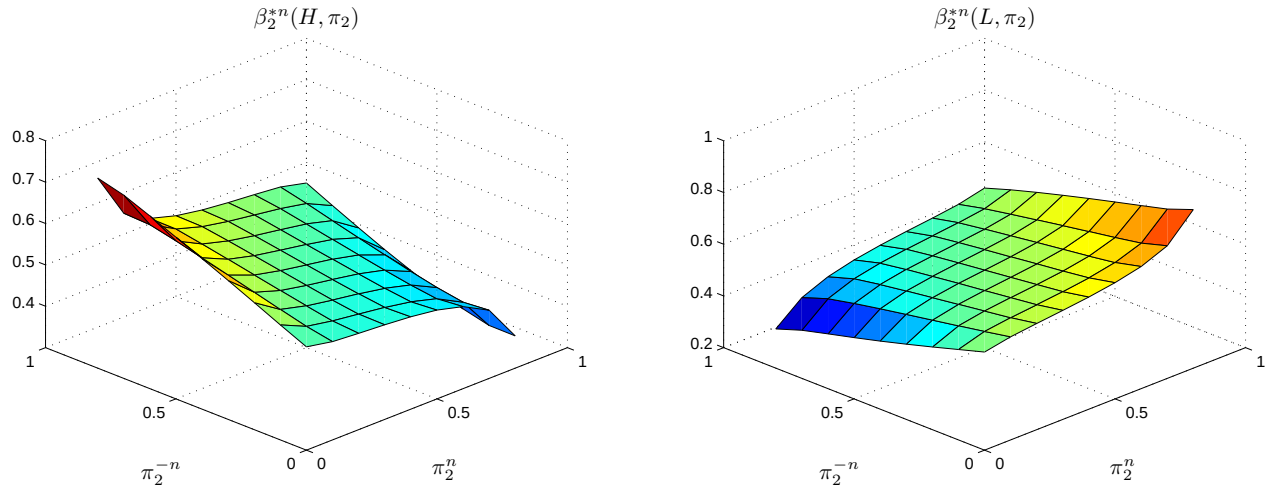

Fig. 2. The strategy $\lambda_2^{*S}(\cdot)$ for the sellers at time $t = 2$.

if for each $\pi_2 \in \Delta(\{H, L\}^2)$ and for each $n = 1, 2$ one the the following conditions is satisfied:

1) $1 > \beta_2^{*n}(H, \pi_2) > 0$ or $1 > \beta_2^{*n}(L, \pi_2) > 0$, and $r_2^{n,\beta_2^*}(\pi_2, h) = r_2^{n,\beta_2^*}(\pi_2, l)$.
2) $\beta_2^{*n}(H, \pi_2) = \beta_2^{*n}(L, \pi_2) = 1$, and $r_2^{n,\beta_2^*}(\pi_2, h) \geq r_2^{n,\beta_2^*}(\pi_2, l)$.
3) $\beta_2^{*n}(H, \pi_2) = \beta_2^{*n}(L, \pi_2) = 0$, and $r_2^{n,\beta_2^*}(\pi_2, h) \leq r_2^{n,\beta_2^*}(\pi_2, l)$.

where $\beta_2^* = \{\beta_2^{*n}(H, \pi_2), \beta_2^{*n}(L, \pi_2), n = 1, 2\}$ are the functions corresponding to $\lambda_2^{*S}(\cdot)$ according to (49) and (50). We numerically compute such a BNE $\lambda_2^{*S}(\cdot)$ along with the functions $\beta_2^{*n}(H, \pi_2)$ and $\beta_2^{*n}(L, \pi_2)$. They are shown in Fig. 2.

Now consider at time 1. Under the CIB strategy profile $(\lambda^S, \lambda^{*B})$ and the CIB update rule ψ^{β_2}, the instantaneous revenue of seller n is

$$p_1^n \mathbb{E}_{\pi_1}^{(\lambda^S, \lambda^{*B})} \left[\sum_{m \in \mathcal{M}} D_1^m(n) \right] = r_1^{n,\beta_1}(\pi_1, p_1^n) \tag{60}$$

where

$$r_1^{n,\beta_1}(\pi_1, p_1^n) := \frac{M}{2(M-1)} p_1^n [y_D(\pi_1^n - \pi_1^{-n}) - p_1^n + \bar{p}_1^{-n}(\pi_1, \beta_1) + (M-1)], \tag{61}$$

and

$$\bar{p}_1^n(\pi_1, \beta_1) := \beta_1^n(\pi_1) h + (1 - \beta_1^n(\pi_1)) l, \tag{62}$$

for $n = 1, 2$. Because of (48), each seller's expected continuation revenue in the analogue of (35) for our example does not depend on $\lambda_1^S(\cdot)$ (see the Appendix for more details). Therefore, the sellers' problem at $t = 1$ is to determine a CIB equilibrium for the game with objectives given by (60). There exists a BNE $\lambda_1^{*S}(\cdot)$ at time $t = 1$ for the game where each seller has the objective given by (60), if for each $\pi_1 \in \Delta(\{H, L\}^2)$ and for each $n = 1, 2$ one the the following conditions is satisfied:

1) $1 > \beta^{*n}(\pi_1) > 0$ and $r_1^{n,\beta_1^*}(\pi_1, h) = r_1^{n,\beta_1^*}(\pi_1, l)$.
2) $\beta^{*n}(\pi_1) = 1$ and $r_1^{n,\beta_1^*}(\pi_1, h) \geq r_1^{n,\beta_1^*}(\pi_1, l)$.
3) $\beta^{*n}(\pi_1) = 0$ and $r_1^{n,\beta_1^*}(\pi_1, h) \leq r_1^{n,\beta_1^*}(\pi_1, l)$.

where $\beta_1^* = \{\beta_1^{*n}(\pi_2), n = 1, 2\}$ are the functions corresponding to $\lambda_1^{*S}(\cdot)$ according (48). We numerically compute such a BNE $\lambda_1^{*S}(\cdot)$ along with the function $\beta_1^{*n}(\pi_1)$. They are shown in Fig. 3. The BNE $\lambda_1^{*S}(\cdot)$ is a solution to the

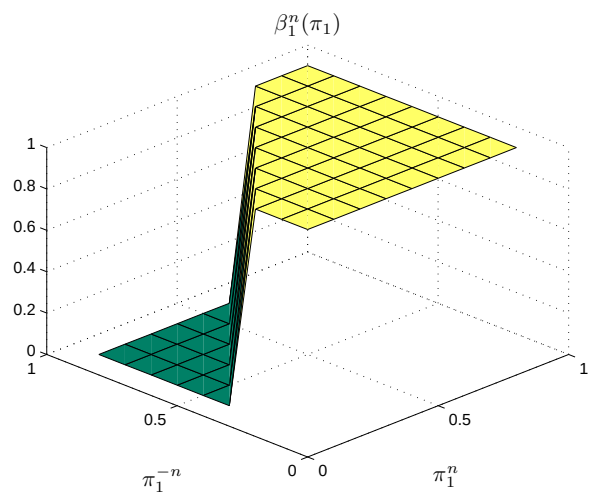

Fig. 3. The strategy $\lambda_1^{*S}(\cdot)$ for the sellers at time $t = 1$.

analogue of (35) in Theorem 2 for our example by the reasons explained above.

From the above calculation we have a CIB strategy $\lambda^* = \{\lambda^{*B}, \lambda^{*S}\}$ and a CIB update rule $\psi^* := \psi^{\beta_2^*}$ such that λ^* solves the dynamic program of Theorem 2, and ψ^* is consistent with λ^*. Then, according to Theorem 2, (λ^*, ψ^*) forms a CIB equilibrium for the dynamic oligopoly game described in this example.

We note that at time $t = 2$, the CIB strategy λ_2^{S*} given by (49) and (50) depends on the private information of the seller at $t = 2$. Consequently, the CIB belief at $t = 2_+$ given

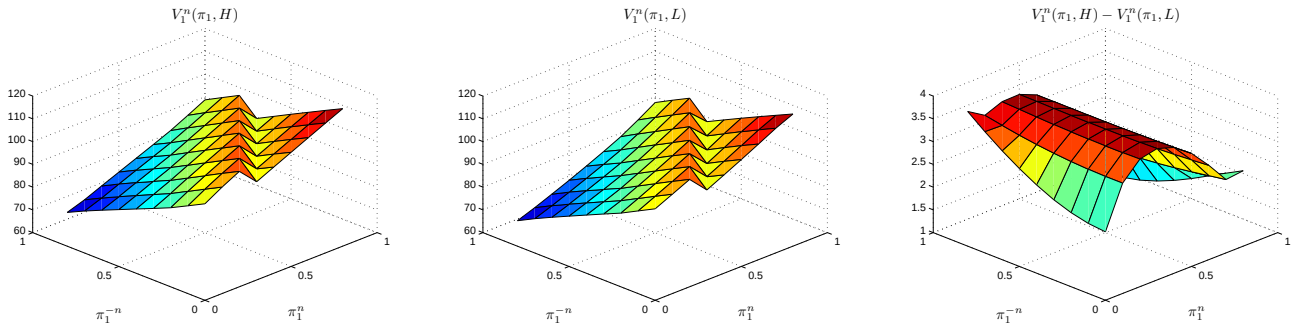

Fig. 4. The expected profit of a seller with state H and that of a seller with state L and their difference.

by (55) and (56) becomes policy-dependent. The policy-dependent model and solution presented in this example highlight the differences between our problem and the ones analyzed in [6], [7].

The expected profit of a seller with state H and that of a seller with state L along with their difference at this equilibrium are shown in Fig. 4.

V. Conclusion

We analyzed a dynamic oligopoly model with private Markovian dynamics. This model gives rise to a stochastic dynamic game with asymmetric information. We used ideas from the common information approach to prove the existence of CIB equilibria, and to obtain a sequential decomposition of the game that leads to a backward induction algorithm for the computation of such equilibria. We illustrated our results with an example. A key feature of our problem is that the CIB beliefs are policy-dependent. As a result, signaling is present in the stochastic dynamic game resulting from our model. This is in contrast to the models of [6], [7] where there is no signaling.

References

[1] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1644–1658, 2013.

[2] D. Fudenberg and J. Tirole, "Game theory. 1991," *Cambridge, Massachusetts*, vol. 393, 1991.

[3] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*, vol. 200. SIAM, 1995.

[4] R. B. Myerson, *Game theory*. Harvard university press, 2013.

[5] J. Filar and K. Vrieze, *Competitive Markov decision processes*. Springer-Verlag New York, Inc., 1996.

[6] A. Nayyar, A. Gupta, C. Langbort, and T. Basar, "Common information based markov perfect equilibria for stochastic games with asymmetric information: Finite games," *IEEE Transactions on Automatic Control*, vol. 59, pp. 555–570, March 2014.

[7] A. Gupta, A. Nayyar, C. Langbort, and T. Basar, "Common information based markov perfect equilibria for linear-gaussian games with asymmetric information," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 3228–3260, 2014.

[8] L. Li and J. Shamma, "Lp formulation of asymmetric zero-sum stochastic games," in *2014 IEEE 53rd Annual Conference on Decision and Control (CDC)*, pp. 7752–7757, 2014.

[9] J. Renault, "The value of markov chain games with lack of information on one side," *Mathematics of Operations Research*, vol. 31, no. 3, pp. 490–512, 2006.

[10] F. Gensbittel and J. Renault, "The value of markov chain games with incomplete information on both sides," *arXiv preprint arXiv:1210.7221*, 2014.

[11] J. Renault, "The value of repeated games with an informed controller," *Mathematics of operations Research*, vol. 37, no. 1, pp. 154–179, 2012.

[12] G. J. Mailath and L. Samuelson, *Repeated games and reputations*, vol. 2. Oxford university press Oxford, 2006.

[13] E. Maskin and J. Tirole, "Markov perfect equilibrium: I. observable actions," *Journal of Economic Theory*, vol. 100, no. 2, pp. 191–219, 2001.

[14] D. Bergemann and J. Välimäki, "Dynamic price competition," *Journal of Economic Theory*, vol. 127, no. 1, pp. 232–263, 2006.

[15] S. Athey, K. Bagwell, and C. Sanchirico, "Collusion and price rigidity," *The Review of Economic Studies*, vol. 71, no. 2, pp. 317–349, 2004.

[16] D. A. Miller, "Robust collusion with private information," *The Review of Economic Studies*, vol. 79, no. 2, pp. 778–811, 2012.

[17] Y. Chen and J. Xie, "Online consumer review: Word-of-mouth as a new element of marketing communication mix," *Management Science*, vol. 54, no. 3, pp. 477–491, 2008.

[18] B. Bickart and R. M. Schindler, "Internet forums as influential sources of consumer information," *Journal of interactive marketing*, vol. 15, no. 3, pp. 31–40, 2001.

[19] J. A. Chevalier and D. Mayzlin, "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research*, vol. 43, no. 3, pp. 345–354, 2006.

[20] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, 2007.

[21] R. Berry, M. L. Honig, and R. Vohra, "Spectrum markets: motivation, challenges, and implications," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 146–155, 2010.

[22] M. Matinmikko, M. Mustonen, D. Roberson, J. Paavola, M. Hoyhtya, S. Yrjola, and J. Roning, "Overview and comparison of recent spectrum sharing approaches in regulation and research: From opportunistic unlicensed access towards licensed shared access," in *2014 IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)*, pp. 92–102, 2014.

[23] O. Ileri, D. Samardzija, and N. B. Mandayam, "Demand responsive pricing and competitive spectrum allocation via a spectrum server," in *2005 First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005.*, pp. 194–202, 2005.

[24] C. Liu and R. A. Berry, "Competition with shared spectrum," in *2014 IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)*, pp. 498–509, 2014.

[25] J. Acharya and R. D. Yates, "Service provider competition and pricing for dynamic spectrum allocation," in *International Conference on Game Theory for Networks, 2009. GameNets' 09.*, pp. 190–198, 2009.

[26] K. R. Liu and B. Wang, *Cognitive radio networking and security: A game-theoretic view*. Cambridge University Press, 2010.

[27] P. Battigalli, "Strategic independence and perfect bayesian equilibria," *Journal of Economic Theory*, vol. 70, no. 1, pp. 201–234, 1996.

[28] P. Kumar and P. Varaiya, *Stochastic Systems: Estimation Identification and Adaptive Control*. Prentice-Hall, Inc., 1986.

APPENDIX

For notational simplicity, we use h_t^n to denote $h_t^{n,S}$ and h_{t+}^c to denote $h_t^{m,B}$ in the proofs.

Proof of Lemma 1. If λ is a CIB strategy profile and ψ is a CIB update rule consistent with λ, we define $g \in \mathcal{G}$ to be a strategy profile where

$$g_t^{n,S}(h_t^n) := \lambda_t^{n,S}(x_t^n, \gamma_\psi(h_t^c)) = \lambda_t^{n,S}(x_t^n, \pi_t^{\gamma_\psi}) \text{ for all } n \in \mathcal{N}, \tag{63}$$

$$g_t^{m,B}(h_{t+}^c) := \lambda_t^{m,B}(p_t, \gamma_\psi(h_{t+}^c)) = \lambda_t^{m,B}(p_t, \pi_{t+}^{\gamma_\psi}) \text{ for all } n \in \mathcal{N}, \tag{64}$$

Then, we recursively define a belief system μ and show by induction that (28) and (29) are satisfied and μ is consistent with g. In order to show the consistency of the belief system, we extend the domain of μ to include histories h_t^c and $h_{t+}^n := (h_t^n, p_t)$ at any time $t \in \mathcal{T}$.

Define $\mu(h_{0+}^c)(\cdot) := 0$ and $\mu(h_{0+}^n)(\cdot) := 0$ for $n \in \mathcal{N}$ at time 0. At time $t = 1$ define, for all $x_1 \in \mathcal{X}^N$,

$$\mu(h_1^c)(x_1) := \mathbb{P}(x_1) \tag{65}$$

$$\mu(h_1^n)(x_1) := \mathbf{1}_{\{x_1^n\}}(h_1^n)\mathbb{P}(x_1^{-n}) \tag{66}$$

for any histories $h_1^n, n \in \mathcal{N}$, and h_1^c. Then, (29) is satisfied at time 0 and (28) is satisfied at time 1, and g is consistent with μ before time 1.

Suppose $\mu(h_t^c)(\cdot)$, $\mu(h_t^n)(\cdot)$, $\mu(h_{(t-1)+}^c)(\cdot)$ and $\mu(h_{(t-1)+}^n)(\cdot)$ are defined and (29) is satisfied at time $(t-1)+$ and (28) is satisfied at time t, and g is consistent with μ before time t (the induction hypothesis).

We proceed to define $\mu(h_{t+1}^n)(\cdot)$, $\mu(h_{t+1}^c)(\cdot)$ and $\mu(h_{t+}^c)(\cdot)$ and prove that (29) is satisfied at time $t+$ and (28) is satisfied at time $t+1$, and g is consistent with μ before time $t+1$.
For any histories h_{t+}^c and $h_{t+}^n, n \in \mathcal{N}$ define the beliefs

$$\mu(h_{t+}^c)(x_{1:t}) := \prod_{k \in \mathcal{N}} \mu(h_{t+}^c)(x_{1:t}^k) \tag{67}$$

$$\mu(h_{t+}^n)(x_{1:t}) := \mathbf{1}_{\{x_{1:t}^n\}}(h_{t+}^n) \prod_{k \neq n} \mu(h_{t+}^c)(x_{1:t}^k) \tag{68}$$

where for any $k \in \mathcal{N}$

$$\mu(h_{t+}^c)(x_{1:t}^k) := \begin{cases} \mu(h_t^c)(x_{1:t}^k)\frac{\gamma_\psi(h_{t+}^c)(x_t^k)}{\gamma_\psi(h_t^c)(x_t^k)}, \\ \quad \text{when } \gamma_\psi(h_t^c)(x_t^k) \neq 0 \\ 0, \text{ when } \gamma_\psi(h_t^c)(x_t^k) = 0. \end{cases} \tag{69}$$

Then, at time $t+$, both sides of (29) are zero when $\gamma_\psi(h_t^c)(x_t) = \prod_{k \in \mathcal{N}} \gamma_\psi(h_t^c)(x_t^k) = 0$. When $\gamma_\psi(h_t^c)(x_t) \neq 0$ we get

$$\begin{aligned} &\mu(h_{t+}^c)(x_t) \\ &= \sum_{x_{1:(t-1)} \in \mathcal{X}^{N(t-1)}} \mu(h_{t+}^c)(x_{1:t}) \\ &= \sum_{x_{1:(t-1)} \in \mathcal{X}^{N(t-1)}} \prod_{k \in \mathcal{N}} \mu(h_t^c)(x_{1:t}^k) \frac{\gamma_\psi(h_{t+}^c)(x_t^k)}{\gamma_\psi(h_t^c)(x_t^k)} \\ &= \prod_{k \in \mathcal{N}} \mu(h_t^c)(x_t^k) \frac{\gamma_\psi(h_{t+}^c)(x_t^k)}{\gamma_\psi(h_t^c)(x_t^k)} \\ &= \prod_{k \in \mathcal{N}} \gamma_\psi(h_{t+}^c)(x_t^k) \end{aligned} \tag{70}$$

where the first and second equalities in (70) follow from (67) and (69), respectively; the last equality in (70) follows from the induction hypothesis for (29) at time t. Therefore, (29) is true at time $t+$.

To show the consistency at time $t+$, we need to show that the Bayes rules, given by

$$\mu(h_{t+}^c)(x_{1:t}) = \frac{\mathbb{P}^g_{\mu(h_t^c)}(x_{1:t}, p_t | h_t^c)}{\sum_{x'_{1:t}} \mathbb{P}^g_{\mu(h_t^c)}(x'_{1:t}, p_t | h_t^c)} \tag{71}$$

$$\mu(h_{t+}^n)(x_{1:t}) = \frac{\mathbb{P}^g_{\mu(h_t^n)}(x_{1:t}, p_t | h_t^n)}{\sum_{x'_{1:t}} \mathbb{P}^g_{\mu(h_t^n)}(x'_{1:t}, p_t | h_t^n)} \tag{72}$$

are true for any histories h_{t+}^c and $h_{t+}^n, n \in \mathcal{N}$ when the above denominators are non-zero. From (67) and (69) we obtain

$$\begin{aligned} &\mu(h_{t+}^c)(x_{1:t}) \\ &= \prod_{k \in \mathcal{N}} \mu(h_{t+}^c)(x_{1:t}^k) \\ &= \prod_{k \in \mathcal{N}} \mu(h_t^c)(x_{1:t}^k) \frac{\gamma_\psi(h_{t+}^c)(x_t^k)}{\gamma_\psi(h_t^c)(x_t^k)} \\ &= \mu(h_t^c)(x_{1:t}) \prod_{k \in \mathcal{N}} \frac{\gamma_\psi(h_{t+}^c)(x_t^k)}{\gamma_\psi(h_t^c)(x_t^k)} \\ &= \mu(h_t^c)(x_{1:t}) \\ &\quad \prod_{k \in \mathcal{N}} \frac{\lambda_t^{k,S}(x_t^k, \gamma_\psi(h_t^c))(p_t^k)}{\sum_{x_t'^k \in \mathcal{X}} \lambda_t^{k,S}(x_t'^k, \gamma_\psi(h_t^c))(p_t^k)\gamma_\psi(h_t^c)(x_t'^k)} \end{aligned} \tag{73}$$

when the denominators are non-zero. The last inequality in (73) is true because of (24) (ψ is consistent with λ).
On the other hand, from the specification of g, the numerator in the right hand side of (71) is equal to

$$\begin{aligned} &\mathbb{P}^g_{\mu(h_t^c)}(x_{1:t}, p_t | h_t^c) \\ &= \mu(h_t^c)(x_{1:t}) \prod_{k \in \mathcal{N}} \lambda_t^{k,S}(x_t^k, \gamma_\psi(h_t^c))(p_t^k). \end{aligned} \tag{74}$$

Substituting (74) back into the right hand side of (71), the right hand side of (71) equals to (73), which is the left hand side of (71). Using similar arguments as in (71), we can show that (72) is true for all $n \in \mathcal{N}$. Therefore, we obtain the consistency of μ at time $t+$.

At time $t+1$ for any histories $h^c_{t+1}, n \in \mathcal{N}$, and h^n_{t+1}, define the belief at time $t+1$ as

$$\mu(h^c_{t+1})(x_{1:t+1}) := \prod_{k\in\mathcal{N}} \mu(h^c_{t+1})(x^k_{1:t+1}) \tag{75}$$

$$\mu(h^n_{t+1})(x_{1:t+1}) := \mathbf{1}_{\{x^n_{1:t+1}\}}(h^n_{t+1}) \prod_{k\neq n} \mu(h^n_{t+1})(x^k_{1:t+1}) \tag{76}$$

where for any $k \in \mathcal{N}$

$$\mu(h^c_{t+1})(x^k_{1:t+1}) := \begin{cases} \mu(h^c_{t+})(x^k_{1:t}) \frac{Q^k(x^k_t, x^k_{t+1}) Q^y(x^k_t, y^k_t)}{\sum_{x'^k_t \in \mathcal{X}} Q^y(x'^k_t, y^k_t)\gamma_\psi(h^c_{t+})(x'^k_t)}, \\ \quad \text{when the above denominator is non-zero,} \\ \frac{1}{|\mathcal{X}|^t}\gamma_\psi(h^c_{t+1})(x^k_{t+1}) \\ \quad \text{when the above denominator is non-zero.} \end{cases} \tag{77}$$

$$\mu(h^n_{t+1})(x^k_{1:t+1}) := \mu(h^c_{t+1})(x^k_{1:t+1}) \text{ for } k \neq n. \tag{78}$$

For any $k \in \mathcal{N}$, when $\sum_{x'^k_t \in \mathcal{X}} Q^y(x'^k_t, y^k_t)\gamma_\psi(h^c_{t+})(x'^k_t) = 0$, we have

$$\begin{aligned} &\mu(h^c_{t+1})(x^k_{t+1}) \\ &= \sum_{x^k_{1:t}\in\mathcal{X}^t} \mu(h^c_{t+1})(x^k_{1:t+1}) \\ &= \sum_{x^k_{1:t}\in\mathcal{X}^t} \frac{1}{|\mathcal{X}|^t}\gamma_\psi(h^c_{t+1})(x^k_{t+1}) \\ &= \gamma_\psi(h^c_{t+1})(x^k_{t+1}); \end{aligned} \tag{79}$$

when $\sum_{x'^k_t \in \mathcal{X}} Q^y(x'^k_t, y^k_t)\gamma_\psi(h^c_{t+})(x'^k_t) \neq 0$, we get

$$\begin{aligned} &\mu(h^c_{t+1})(x^k_{t+1}) \\ &= \sum_{x^k_{1:t}\in\mathcal{X}^t} \mu(h^c_{t+1})(x^k_{1:t+1}) \\ &= \sum_{x^k_{1:t}\in\mathcal{X}^t} \mu(h^c_{t+})(x^k_{1:t}) \frac{Q^k(x^k_t, x^k_{t+1}) Q^y(x^k_t, y^k_t)}{\sum_{x'^k_t \in \mathcal{X}} Q^y(x'^k_t, y^k_t)\gamma_\psi(h^c_{t+})(x'^k_t)} \\ &= \sum_{x^k_t\in\mathcal{X}} \mu(h^c_{t+})(x^k_t) \frac{Q^k(x^k_t, x^k_{t+1}) Q^y(x^k_t, y^k_t)}{\sum_{x'^k_t \in \mathcal{X}} Q^y(x'^k_t, y^k_t)\gamma_\psi(h^c_{t+})(x'^k_t)} \\ &= \sum_{x^k_t\in\mathcal{X}} \gamma_\psi(h^c_{t+})(x^k_t) \frac{Q^k(x^k_t, x^k_{t+1}) Q^y(x^k_t, y^k_t)}{\sum_{x'^k_t \in \mathcal{X}} Q^y(x'^k_t, y^k_t)\gamma_\psi(h^c_{t+})(x'^k_t)} \\ &= \gamma_\psi(h^c_{t+1})(x^k_{t+1}). \end{aligned} \tag{80}$$

where the forth equality in (80) follows from the induction hypothesis for (28) at time $t+$ and the last equality in (80) is true because of (25) (ψ is consistent with λ). Then, for any $n \in \mathcal{N}$, from (76), (79) and (80) we obtain

$$\begin{aligned} &\mu(h^n_{t+1})(x_{t+1}) \\ &= \sum_{x_{1:t}} \mathbf{1}_{\{x^n_{1:t+1}\}}(h^n_{t+1}) \prod_{k\neq n} \mu(h^c_{t+1})(x^k_{1:t+1}) \\ &= \mathbf{1}_{\{x^n_{t+1}\}}(h^n_{t+1}) \prod_{k\neq n} \mu(h^c_{t+1})(x^k_{t+1}) \\ &= \mathbf{1}_{\{x^n_{t+1}\}}(h^n_{t+1}) \prod_{k\neq n} \gamma_\psi(h^c_{t+1})(x^k_{t+1}). \end{aligned} \tag{81}$$

Therefore, (28) is true at time $t+1$.

To show the consistency at time $t+1$, we need to show that the Bayes rules, given by

$$\begin{aligned} &\mu(h^c_{t+1})(x_{1:t+1}) \\ &= \frac{\mathbb{P}^g_{\mu(h^c_{t+})}(x_{1:t+1}, d_t, y_t | h^c_{t+})}{\sum_{x'_{1:t+1}} \mathbb{P}^g_{\mu(h^c_{t+})}(x'_{1:t+1}, d_t, y_t | h^c_{t+})} \end{aligned} \tag{82}$$

$$\begin{aligned} &\mu(h^n_{t+1})(x_{1:t+1}) \\ &= \frac{\mathbf{1}_{\{x^n_{t+1}\}}(h^n_{t+1}) \mathbb{P}^g_{\mu(h^n_{t+})}(x_{1:t+1}, d_t, y_t | h^n_{t+})}{\sum_{x'_{1:t+1}} \mathbf{1}_{\{x^n_{t+1}\}}(x'_{1:t+1}) \mathbb{P}^g_{\mu(h^n_{t+})}(x'_{1:t+1}, d_t, y_t | h^n_{t+})} \end{aligned} \tag{83}$$

are true for any histories h^c_{t+1} and $h^n_{t+1}, n \in \mathcal{N}$ when the above denominators are non-zero.
From (75) and (77) we obtain

$$\begin{aligned} &\mu(h^c_{t+1})(x_{1:t+1}) \\ &= \prod_{k\in\mathcal{N}} \mu(h^c_{t+})(x^k_{1:t}) \frac{Q^k(x^k_t, x^k_{t+1}) Q^y(x^k_t, y^k_t)}{\sum_{x'^k_t \in \mathcal{X}} Q^y(x'^k_t, y^k_t)\gamma_\psi(h^c_{t+})(x'^k_t)} \\ &= \mu(h^c_{t+})(x_{1:t}) \\ &\quad \prod_{k\in\mathcal{N}} \frac{Q^k(x^k_t, x^k_{t+1}) Q^y(x^k_t, y^k_t)}{\sum_{x'^k_t \in \mathcal{X}} Q^y(x'^k_t, y^k_t)\gamma_\psi(h^c_{t+})(x'^k_t)} \\ &= \mu(h^c_{t+})(x_{1:t}) \\ &\quad \prod_{k\in\mathcal{N}} \frac{Q^k(x^k_t, x^k_{t+1}) Q^y(x^k_t, y^k_t)}{\sum_{x'^k_t \in \mathcal{X}} Q^y(x'^k_t, y^k_t)\mu(h^c_{t+})(x'^k_t)} \end{aligned} \tag{84}$$

when the above denominator is non-zero; the last equality in (84) follows from the induction hypothesis for (29) at time $t+$.
On the other hand, using the dynamics of states and the specification of g, the numerator in the right hand side of (82) is equal to

$$\begin{aligned} &\mathbb{P}^g_{\mu(h^c_{t+})}(x_{1:t+1}, d_t, y_t | h^c_{t+}) \\ &= \mu(h^c_{t+})(x_{1:t}) \prod_{m\in\mathcal{M}} g^{m,B}(h^c_{t+}) \\ &\quad \prod_{k\in\mathcal{N}} Q^k(x^k_t, x^k_{t+1}) Q^y(x^k_t, y^k_t). \end{aligned} \tag{85}$$

Substituting (85) back into the right hand side of (82), the right hand side of (82) equals to (84), which is the left hand side of (82). Using similar arguments as in (82), we can show that (83) is true for all $n \in \mathcal{N}$. Therefore, we obtain the consistency of μ at time $t+1$.

□

Proof of Lemma 2. Due to space limit, we only prove the case for seller n; the buyer's case can be proved using arguments similar to the seller's case. To simply the notation, we use γ to denote the CIB belief γ_ψ constructed from ψ.

Let $(g, \mu) = f(\lambda, \psi)$. Suppose every buyer m uses the strategy $g^{m,B}$ and every seller $k \neq n$ uses the strategy $g^{k,S}$. Since the strategy of every buyer and every seller $k \neq n$ is fixed, the best response of seller n is the solution to the

following stochastic control problem.

$$\max_{\tilde{g}^{n,S}\in\mathcal{G}^{n,S}} \mathbb{E}_{\mu}^{\tilde{g}^{n,S}}\left[\sum_{t=1}^{T}\phi_t^{n,S}(P_t,D_t)\right]. \tag{86}$$

In the following, we show that the stochastic control problem (86) is equivalent to a Markov Decision Process (MDP) with state process $\{(X_t^n,\Pi_t^\gamma), t\in\mathcal{T}\}$ and action process $\{P_t^n, t\in\mathcal{T}\}$ of seller n.

Since $P_t^k, k\neq n$ satisfies (26), the distribution of P_t^k only depends on $\{X_t^k,\Pi_t^\gamma\}$. Each D_t^m satisfies (27), then the distribution of D_t^m depends on $\{P_t,\Pi_{t+}^\gamma\}$. Note that $\Pi_{t+}^\gamma=\psi_t(\Pi_t^\gamma,P_t)$ from (18). Consequently, the distribution of D_t^m only depends on $\{P_t^n,X_t^{-n},\Pi_t^\gamma\}$.
Therefore, the distribution of (P_t,D_t) only depends on $\{P_t^n,X_t^{-n},\Pi_t^\gamma\}$. Then,

$$\begin{aligned}&\mathbb{E}_{\mu}^{\tilde{g}^{n,S}}\left[\phi_t^{n,S}(P_t,D_t)|P_t^n,X_t^{-n},\Pi_t^\gamma,H_t^n\right]\\ =&\mathbb{E}\left[\phi_t^{n,S}(P_t,D_t)|P_t^n,X_t^{-n},\Pi_t^\gamma\right]\\ =:&\bar{\phi}_t^{n,S}(P_t^n,X_t^{-n},\Pi_t^\gamma).\end{aligned} \tag{87}$$

For any realization $h_t^n=(h_t^c,x_{1:t}^n)\in\mathcal{H}_t^{n,S}$ and $p_t^n\in\mathcal{P}$, letting $\pi_t^\gamma:=\gamma(h_t^c)$. From (87) we obtain

$$\begin{aligned}&\mathbb{E}_{\mu}^{\tilde{g}^{n,S}}\left[\phi_t^{n,S}(P_t,D_t)|h_t^n,p_t^n\right]\\ =&\mathbb{E}_{\mu}^{\tilde{g}^{n,S}}\left[\mathbb{E}\left[\phi_t^{n,S}(P_t,D_t)|P_t^n,X_t^{-n},\Pi_t^\gamma,H_t^n\right]|h_t^n,p_t^n\right]\\ =&\mathbb{E}_{\mu}\left[\bar{\phi}_t^{n,S}(p_t^n,X_t^{-n},\pi_t^\gamma)|h_t^c,x_{1:t}^n,p_t^n\right]\\ \overset{(a)}{=}&\mathbb{E}_{\mu}\left[\bar{\phi}_t^{n,S}(p_t^n,X_t^{-n},\pi_t^\gamma)|h_t^c\right]\\ =:&\tilde{\phi}_t^{n,S}(\pi_t^\gamma,p_t^n).\end{aligned} \tag{88}$$

Equation (a) in (88) is true because from (28), X_t^{-n} and X_t^n are independent under μ. From (88), the objective function (86) of seller n is equivalent to

$$\max_{\tilde{g}^{n,S}\in\mathcal{G}^{n,S}} \mathbb{E}_{\mu}^{\tilde{g}^{n,S}}\left[\sum_{t=1}^{T}\tilde{\phi}_t^{n,S}(\Pi_t^\gamma,P_t^n)\right]. \tag{89}$$

From the private state dynamics, the distribution of X_{t+1}^n depends only on X_t^n. From (16) and (18) we get

$$\begin{aligned}\Pi_{t+1}^\gamma=&\psi_{t+}(\Pi_{t+}^\gamma,Y_t)\\ =&\psi_{t+}(\psi_t(\Pi_t^\gamma,P_t),Y_t).\end{aligned} \tag{90}$$

Note that the distribution of P_t^k only depends on $\{X_t^k,\Pi_t^\gamma\}$; the distribution of Y_t depends only on (X_t^n,X_t^{-n}). Therefore, the distribution of $(X_{t+1}^n,X_{t+1},\Pi_{t+1}^\gamma)$ depends only on $\{X_t^n,X_t^{-n},P_t^n,\Pi_t^\gamma\}$. Then, for any realizations $x_{t+1}^n\in\mathcal{X}$, $\pi_{t+1}^\gamma\in\Delta(\mathcal{X})$, $h_t^n=(h_t^c,x_{1:t}^n)\in\mathcal{H}_t^{n,S}$ and $p_t^n\in\mathcal{P}$, we obtain

$$\begin{aligned}&\mathbb{P}_{\mu}^{\tilde{g}^{n,S}}(x_{t+1}^n,\pi_{t+1}^\gamma|h_t^n,p_t^n)\\ =&\sum_{x_t^{-n}}\mathbb{P}_{\mu}^{\tilde{g}^{n,S}}(x_{t+1}^n,\pi_{t+1}^\gamma|x_t^{-n},h_t^n,p_t^n)\mathbb{P}_{\mu}^{\tilde{g}^{n,S}}(x_t^{-n}|h_t^n,p_t^n)\\ =&\sum_{x_t^{-n}}\mathbb{P}(x_{t+1}^n,\pi_{t+1}^\gamma|\pi_t^\gamma,x_t,p_t^n)\mu(h_t^n)(x_t^{-n})\\ =&\sum_{x_t^{-n}}\mathbb{P}(x_{t+1}^n,\pi_{t+1}^\gamma|\pi_t^\gamma,x_t,p_t^n)\prod_{k\neq n}\pi_t^{k,\gamma}(x_t^k)\\ =&\mathbb{P}_{\mu}^{\tilde{g}^{n,S}}(x_{t+1}^n,\pi_{t+1}^\gamma|x_t^n,\pi_t^\gamma,p_t^n).\end{aligned} \tag{91}$$

The third equality in (91) follows from Lemma 1, and the last equality follows from the same arguments as the first through third equalities. Equation (91) shows that the process $\{(X_t^n,\Pi_t^\gamma), t\in\mathcal{T}\}$ is a controlled Markov Chain with respect to the action process $\{P_t^n, t\in\mathcal{T}\}$ for seller n. Therefore, (89) is a MDP. From the theory of MDP (see [28]), there is a best response $\tilde{g}^{n,S}$, under the belief system μ, of seller n such that

$$\tilde{g}^{n,S}(H_t^n)=\tilde{\lambda}^{n,S}(X_t^n,\Pi_t^\gamma)=\tilde{\lambda}^{n,S}(X_t^n,\gamma(H_t^c)) \tag{92}$$

for some function $\tilde{\lambda}^{n,S}(\cdot)$ at each time t. This completes the proof of Lemma 2 for the sellers.

□

Proof of Theorem 2. If a pair (λ^*,ψ^*) is a CIB equilibrium, let $(g^*,\mu^*)=f(\lambda^*,\psi^*)$ be the PBE constructed from (λ^*,ψ^*). From the definition of the functions $V^*:=V(\lambda^*,\psi^*)$, $V_t^{*n,S}(\gamma(h_t^c),x_t^n)$ is the expected reward to go from time t on, under μ^*, for seller n at h_t^n under the strategy profile g^*. If (35) is not true, then there exists another strategy of seller n that achieves reward higher than $V_t^{*n,S}(\gamma(h_t^c),x_t^n)$. This contradicts the fact that (g^*,μ^*) is a PBE. Similar arguments show that (36) should also hold.

On the other hand, if λ^* solves the dynamic program (35)-(36) of the functions $V^*:=V(\lambda^*,\psi^*)$, let $(g^*,\mu^*)=f(\lambda^*,\psi^*)$. Then μ^* is consistent with g^* from Lemma 1.

Due to space limit, we only prove sequential rationality for seller n. Sequential rationality for the buyers can be obtained using similar arguments.

If every buyer m uses the strategy $g^{*m,B}$ and every seller $k\neq n$ uses the strategy $g^{*k,S}$, from Lemma 2 we know that there is a best response $\tilde{g}^{n,S}$, under the belief system μ^*, of seller n such that

$$\tilde{g}_t^{n,S}(h_t^n)=\tilde{\lambda}_t^{n,S}(x_t^n,\pi_t^{\gamma^*}) \tag{93}$$

for some CIB strategy $\tilde{\lambda}_t^{n,S}(\cdot)$ at each time t. Define a CIB strategy profile $\tilde{\lambda}:=(\tilde{\lambda}^{n,S},\lambda^{*-n,S},\lambda^{*B})$. From the definition of the functions $\tilde{V}:=V(\tilde{\lambda},\psi^*)$, $\tilde{V}_t^{n,S}(\gamma(h_t^c),x_t^n)$ is the expected reward to go from time t on, under μ^*, for seller n at h_t^n under the strategy profile $\tilde{g}$. Since $\tilde{g}^{n,S}$ is a best response, $\tilde{V}_t^{n,S}(\gamma(h_t^c),x_t^n)$ gives seller n the maximum expected reward to go from time t on under μ^*. However, since $\lambda_t^{*n,S}(x_t^n,\pi_t^{\gamma^*})$ is one of the optimal solution of (35) at each time t, we can show by induction that

$$V_t^{*n,S}(\gamma(h_t^c),x_t^n)\geq\tilde{V}_t^{n,S}(\gamma(h_t^c),x_t^n). \tag{94}$$

Therefore, (94) implies that, at any time t, $g^{*n,S}_{t:T}$ gives seller n the maximum expected reward to go from time t on under μ^*. This complete the proof that (g^*, μ^*) is a PBE. As a result, the pair (λ^*, γ^*) forms a CIB equilibrium of the dynamic oligopoly game described in II. □

Proof of Theorem 1. Consider a CIB update rule ψ^* given by

$$\psi^{*n}_t(\pi_t, p^n_t) = \pi^n_t \tag{95}$$

$$\psi^{*n}_{t+}(\pi_{t+}, y^n_t)(x^n_{t+1}) = \begin{cases} \frac{\sum_{x^n_t} Q^n(x^n_t, x^n_{t+1}) Q^y(x^n_t, y^n_t) \pi^n_{t+}(x^n_t)}{\sum_{x'^n_t \in \mathcal{X}} Q^y(x'^n_t, y^n_t) \pi^n_{t+}(x'^n_t)}, \\ \quad \text{when the above denominator is non-zero.} \\ \mathbf{1}_{\{x^n_{t+1}=1\}}, \text{ otherwise} \end{cases} \tag{96}$$

Based on ψ^*, we solve the dynamic program (36)-(35) to get a CIB strategy profile λ^* and show that (λ^*, ψ^*) forms a CIB equilibrium.

To solve the dynamic program, first consider buyers' strategies. Suppose the users use a CIB strategy profile λ. For the functions $V(\lambda, \psi^*)$, since Y_t is independent of D_t, the optimization problem (36) becomes

$$\begin{aligned} &\arg\max_{d^m_t \in \mathcal{D}} \left\{ \mathbb{E}^{\lambda^{-m,B}_t}_{\pi_{t+}} \left[\phi^{m,B}_t(Y_t, p_t, (d^m_t, D^{-m}_t)) \right.\right. \\ &\quad \left.\left. + V^{m,B}_{t+1}(\psi^*_{t+}(\pi_{t+}, Y_t)) | p_t, \pi_{t+} \right] \right\} \\ =&\arg\max_{d^m_t \in \mathcal{D}} \left\{ \mathbb{E}^{\lambda^{-m,B}_t}_{\pi_{t+}} \left[\phi^{m,B}_t(Y_t, p_t, (d^m_t, D^{-m}_t)) | p_t, \pi_{t+} \right] \right\} \\ &+ \mathbb{E}_{\pi_{t+}} \left[V^{m,B}_{t+1}(\psi^*_{t+}(\pi_{t+}, Y_t)) | p_t, \pi_{t+} \right]. \end{aligned} \tag{97}$$

Consequently, λ^B_t solves (36) if for any $p_t \in \mathcal{P}$ and $\pi_{t+} \in \Delta(\mathcal{X}^N)$, $\{\lambda^{m,B}_t(p_t, \pi_{t+}), m \in \mathcal{M}\}$ is a Bayesian Nash equilibrium of the game $G_{t+}(p_t, \pi_{t+})$ where each agent m has the following objective.

$$\arg\max_{d^m_t \in \mathcal{D}} \left\{ \mathbb{E}_{\pi_{t+}} \left[\phi^{m,B}_t(Y_t, p_t, (d^m_t, d^{-m}_t)) | p_t, \pi_{t+} \right] \right\} \tag{98}$$

Since $\mathcal{D}$ is a finite set, $G_{t+}(p_t, \pi_{t+})$ has at least one equilibrium. Let $\{\lambda^{*m,B}_t(p_t, \pi_{t+}), m \in \mathcal{M}\}$ denote an equilibrium of $G_{t+}(p_t, \pi_{t+})$ for $t \in \mathcal{T}$. Then λ^{*B}_t solves (36) for any time $t \in \mathcal{T}$.

Now suppose the buyers use the strategy λ^{*B} and the sellers use some strategy λ^S. For the functions $V((\lambda^S, \lambda^{*B}), \psi^*)$. Note that $\psi^{*n}_t(\pi_t, p^n_t) = \pi^n_t$. Therefore, the optimization problem (35) becomes

$$\begin{aligned} &\arg\max_{p^n_t \in \mathcal{D}} \left\{ \mathbb{E}^{\lambda^{*-n,S}_t}_{\pi^n_t} \left[V^{*n,S}_{t+}(\right.\right. \\ &\quad \left.\left. \psi^*_t(\pi_t, (p^n_t, P^{-n}_t)), (p^n_t, P^{-n}_t), x^n_t) | \pi_t, x^n_t \right] \right\} \\ =&\arg\max_{p^n_t \in \mathcal{D}} \left\{ \mathbb{E}^{\lambda^{-n,S}_t}_{\pi^n_t} \left[V^{n,S}_{t+}(\pi_t, z_t, (p^n_t, P^{-n}_t), x^n_t) | \pi_t, x^n_t \right] \right\}. \end{aligned} \tag{99}$$

The function $V^{n,S}_{t+}(\cdot)$ can be computed by

$$\begin{aligned} &V^{n,S}_{t+}(\pi_t, p_t, x^n_t) \\ =&\mathbb{E}^{\lambda^{*B}_t}_{\pi^n_t} \left[\phi^{n,S}_t(p_t, D_t) \right. \\ &\quad \left. + V^{n,S}_{t+1}(\psi_t(\pi_t, Y_t), X^n_{t+1}) | \pi_t, p_t, x^n_t \right] \\ =& \sum_{d_t \in \mathcal{D}^M} \phi^{n,S}_t(p_t, d_t) \prod_{m \in \mathcal{M}} \lambda^{*m,B}_t(p_t, \pi_t)(d^m_t) \\ &+ \mathbb{E}_{\pi_{t+}} \left[V^{n,S}_{t+1}(\psi_{t+}(\pi_t, Y_t), X^n_{t+1}) | \pi_t, x^n_t \right] \\ =&\hat{\phi}^{n,S}_t(p_t, \pi_t) + \hat{V}^{n,S}_{t+1}(\pi_t, x^n_t) \end{aligned} \tag{100}$$

where

$$\begin{aligned} &\hat{\phi}^{n,S}_t(p_t, \pi_t) \\ :=& \sum_{d_t \in \mathcal{D}^M} \phi^{n,S}_t(p_t, d_t) \prod_{m \in \mathcal{M}} \lambda^{*m,B}_t(p_t, \pi_t)(d^m_t) \end{aligned} \tag{101}$$

$$\begin{aligned} &\hat{V}^{n,S}_{t+1}(\pi_t, x^n_t) \\ :=&\mathbb{E}_{\pi_{t+}} \left[V^{n,S}_{t+1}(\psi_{t+}(\pi_t, Y_t), X^n_{t+1}) | \pi_t, x^n_t \right]. \end{aligned} \tag{102}$$

Putting (100) into (99), the optimization problem (35) becomes

$$\begin{aligned} &\arg\max_{p^n_t \in \mathcal{D}} \left\{ \mathbb{E}^{\lambda^{-n,S}_t}_{\pi^n_t} \left[\hat{\phi}^{n,S}_t((p^n_t, P^{-n}_t), \pi_t) \right.\right. \\ &\quad \left.\left. + \hat{V}^{n,S}_{t+1}(\pi_t, x^n_t) | x^n_t, \pi_t \right] \right\} \\ =&\arg\max_{p^n_t \in \mathcal{D}} \left\{ \mathbb{E}^{\lambda^{-n,S}_t}_{\pi^n_t} \left[\hat{\phi}^{n,S}_t((p^n_t, P^{-n}_t), \pi_t) | z_t, \pi_t \right] \right\} \\ &+ \hat{V}^{n,S}_{t+1}(\pi_t, x^n_t). \end{aligned} \tag{103}$$

Consequently, λ^S_t solves (35) if $\lambda^{n,S}_t(x^n_t, \pi_t) = \hat{\lambda}^{n,S}_t(\pi_t)$ and for any $\pi_t \in \Delta(\mathcal{X}^N)$, $\{\hat{\lambda}^{n,S}_t(\pi_t), n \in \mathcal{N}\}$ is a Nash equilibrium of the game $G_t(\pi_t)$ where each user n has the following objective.

$$\arg\max_{p^n_t \in \mathcal{P}} \left\{ \hat{\phi}^{n,S}_t((p^n_t, p^{-n}_t), \pi_t)) \right\}. \tag{104}$$

Since $\mathcal{P}$ is a finite set, $G_t(\pi_t)$ has at least one equilibrium. Let $\{\hat{\lambda}^{*n,S}_t(\pi_t), n \in \mathcal{N}\}$ denote an equilibrium of $G_t(\pi_t)$ for $t \in \mathcal{T}$. Let $\{\lambda^{*n,S}_t(x^n_t, \pi_t) := \hat{\lambda}^{*n,S}_t(\pi_t), n \in \mathcal{N}\}$. Then λ^{*S}_t solves (35) for any time $t \in \mathcal{T}$.

From the above calculation we have a CIB strategy profile $\lambda^* = \{\lambda^{*B}, \lambda^{*S}\}$ and a CIB update rule ψ^* such that λ^* solves the dynamic program of the functions $V(\lambda^*, \psi^*)$. Furthermore, ψ^* satisfies (25) by its definition (96). Note that $\lambda^{*n,S}_t(x^n_t, \pi_t) = \hat{\lambda}^{*n,S}_t(\pi_t)$. Therefore,

$$\begin{aligned} &\frac{\lambda^{*n,S}_t(x^n_t, \pi_t)(p^n_t)\pi^n_t(x^n_t)}{\sum_{x'^n_t \in \mathcal{X}} \lambda^{*n,S}_t(x'^n_t, \pi_t)(p^n_t)\pi^n_t(x'^n_t)} \\ =&\frac{\hat{\lambda}^{*n,S}_t(\pi_t)(p^n_t)\pi^n_t(x^n_t)}{\sum_{x'^n_t \in \mathcal{X}} \hat{\lambda}^{*n,S}_t(\pi_t)(p^n_t)\pi^n_t(x'^n_t)} \\ =&\pi^n_t(x^n_t) \\ =&\psi^{*n}_t(\pi_t, p^n_t)(x^n_t) \end{aligned} \tag{105}$$

Then, ψ^* also satisfies (24). As a result, ψ^* is consistent with λ^*. According to Theorem 2, (λ^*, ψ^*) forms a CIB equilibrium. $\square$